GenCore version 5.1.6


OM protein - protein search, using sw model

Run on: July 2, 2003, 19:20:50 ; Search time 23 Seconds
(without alignments)
1609.901 Million cell updates/sec

Title: US-09-849-869A-16
Perfect score: 1661
Sequence: 1 MDPTISTLDTELTPINGTEE.....EGGGQLPEEILELSGSRLEQ 322

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 440863 seqs, 114992915 residues

Total number of hits satisfying chosen parameters: 440863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 1661 | 100.0 | 322 | 9 | US-10-183-116-16 | Sequence 16, Appl |
| 2 | 1661 | 100.0 | 322 | 9 | US-10-079-384-4 | Sequence 4, Appli |
| 3 | 1383 | 83.3 | 322 | 9 | US-09-995-225-20 | Sequence 20, Appl |
| 4 | 1383 | 83.3 | 322 | 9 | US-10-183-116-31 | Sequence 31, Appl |
| 5 | 1383 | 83.3 | 322 | 9 | US-10-225-567A-674 | Sequence 674, App |
| 6 | 1334 | 80.3 | 314 | 9 | US-10-219-834-79 | Sequence 79, Appl |
| 7 | 1294 | 77.9 | 322 | 9 | US-09-995-225-18 | Sequence 18, Appl |
| 8 | 1294 | 77.9 | 322 | 9 | US-10-183-116-33 | Sequence 33, Appl |
| 9 | 1294 | 77.9 | 322 | 9 | US-10-225-567A-689 | Sequence 689, App |
| 10 | 1212 | 73.0 | 328 | 9 | US-10-219-834-20 | Sequence 20, Appl |
| 11 | 1038.5 | 62.5 | 330 | 9 | US-10-183-116-18 | Sequence 18, Appl |
| 12 | 1038.5 | 62.5 | 330 | 9 | US-10-079-384-12 | Sequence 12, Appl |
| 13 | 1038.5 | 62.5 | 330 | 9 | US-10-225-567A-649 | Sequence 649, App |
| 14 | 1038.5 | 62.5 | 330 | 10 | US-09-826-508-12 | Sequence 12, Appl |
| 15 | 1038.5 | 62.5 | 330 | 10 | US-09-750-373-24 | Sequence 24, Appl |
| 16 | 1038.5 | 62.5 | 330 | 10 | US-09-920-068A-4 | Sequence 4, Appli |
| 17 | 1032 | 62.1 | 330 | 9 | US-09-800-321A-31 | Sequence 31, Appl |
| 18 | 1013.5 | 61.0 | 330 | 9 | US-10-012-140-8 | Sequence 8, Appli |
| 19 | 984 | 59.2 | 324 | 9 | US-09-782-974C-58 | Sequence 58, Appl |
| 20 | 821 | 49.4 | 338 | 9 | US-10-183-116-41 | Sequence 41, Appl |
| 21 | 779 | 46.9 | 273 | 9 | US-10-183-116-85 | Sequence 85, Appl |
| 22 | 773.5 | 46.6 | 325 | 10 | US-09-920-068A-3 | Sequence 3, Appli |
| 23 | 769.5 | 46.3 | 338 | 9 | US-10-183-116-39 | Sequence 39, Appl |
| 24 | 764.5 | 46.0 | 303 | 9 | US-10-183-116-53 | Sequence 53, Appl |
| 25 | 756.5 | 45.5 | 304 | 9 | US-10-183-116-2 | Sequence 2, Appli |
| 26 | 748.5 | 45.1 | 305 | 9 | US-10-183-116-57 | Sequence 57, Appl |
| 27 | 745.5 | 44.9 | 270 | 9 | US-10-183-116-87 | Sequence 87, Appl |
| 28 | 735 | 44.3 | 302 | 9 | US-10-183-116-6 | Sequence 6, Appli |
| 29 | 731.5 | 44.0 | 301 | 9 | US-10-183-116-23 | Sequence 23, Appl |
| 30 | 731.5 | 44.0 | 321 | 9 | US-10-183-116-45 | Sequence 45, Appl |
| 31 | 725 | 43.6 | 176 | 9 | US-10-116-252-11 | Sequence 11, Appl |
| 32 | 716.5 | 43.1 | 305 | 9 | US-10-183-116-25 | Sequence 25, Appl |
| 33 | 716.5 | 43.1 | 305 | 9 | US-10-183-116-63 | Sequence 63, Appl |
| 34 | 699.5 | 42.1 | 305 | 9 | US-10-183-116-27 | Sequence 27, Appl |
| 35 | 695 | 41.8 | 331 | 10 | US-09-920-068A-1 | Sequence 1, Appli |
| 36 | 692.5 | 41.7 | 322 | 9 | US-10-183-116-47 | Sequence 47, Appl |
| 37 | 687.5 | 41.4 | 268 | 9 | US-10-183-116-79 | Sequence 79, Appl |
| 38 | 682.5 | 41.1 | 305 | 9 | US-10-183-116-12 | Sequence 12, Appl |
| 39 | 680.5 | 41.0 | 305 | 9 | US-10-183-116-4 | Sequence 4, Appli |
| 40 | 679 | 40.9 | 300 | 9 | US-10-183-116-65 | Sequence 65, Appl |
| 41 | 678.5 | 40.8 | 303 | 9 | US-10-183-116-67 | Sequence 67, Appl |
| 42 | 675 | 40.6 | 287 | 9 | US-10-183-116-21 | Sequence 21, Appl |
| 43 | 672 | 40.5 | 304 | 9 | US-10-183-116-73 | Sequence 73, Appl |
| 44 | 671.5 | 40.4 | 282 | 9 | US-10-183-116-55 | Sequence 55, Appl |
| 45 | 667.5 | 40.2 | 305 | 9 | US-10-183-116-59 | Sequence 59, Appl |

ALIGNMENTS

RESULT 1
US-10-183-116-16
; Sequence 16, Application US/10183116
; Publication No. US20030092035A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David J.
; APPLICANT: Dong, Xinzhong
; APPLICANT: Zylka, Mark
; APPLICANT: Simon, Melvin
; APPLICANT: Han, Sang-kyou
; TITLE OF INVENTION: PAIN SIGNALING MOLECULES
; FILE REFERENCE: CALTE.4C1CP1
; CURRENT APPLICATION NUMBER: US/10/183,116
; CURRENT FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: US 60/222,344
; PRIOR FILING DATE: 2000-08-01
; PRIOR APPLICATION NUMBER: US 60/202,027
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 09/704,707
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: US 60/285,493
; PRIOR FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: US 09/849,869
; PRIOR FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 322
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-183-116-16

Query Match 100.0%; Score 1661; DB 9; Length 322;
Best Local Similarity 100.0%; Pred. No. 4.9e-142;
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDPTISTLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNA 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MDPTISTLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNA 60

Qy 61 FSIYILNLAAADFLFLSGRLIYSLLSFISIPHTISKILYPVMMFSYFAGLSFLSAVSTER 120
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
Db          61 FSIYILNLAAADFLFLSGRLIYSLLSFISIPHTISKILYPVMMFSYFAGLSFLSAVSTER 120

Qy         121 CLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQTSDFIT 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 CLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQTSDFIT 180

Qy         181 VAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGIQFFLFL 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 VAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGIQFFLFL 240

Qy         241 WIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASE 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 WIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASE 300

Qy         301 VDEGGGQLPEEILELSGSRLEQ 322
               ||||||||||||||||||||||
Db         301 VDEGGGQLPEEILELSGSRLEQ 322


RESULT 2
US-10-079-384-4
; Sequence 4, Application US/10079384
; Publication No. US20030108986A1
; GENERAL INFORMATION:
;  APPLICANT: Communi, Didier
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS COMPRISING G-PROTEIN COUPLED RECEPTORS
;  FILE REFERENCE: 9409/2132
;  CURRENT APPLICATION NUMBER: US/10/079,384
;  CURRENT FILING DATE:  2002-02-20
;  PRIOR APPLICATION NUMBER: US 09/885,453
;  PRIOR FILING DATE: 2001-06-20
;  NUMBER OF SEQ ID NOS: 50
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
;   LENGTH: 322
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-079-384-4

  Query Match           100.0%;  Score 1661;  DB 9;  Length 322;
  Best Local Similarity  100.0%;  Pred. No. 4.9e-142;
  Matches  322;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy           1 MDPTISTLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNA 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 MDPTISTLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNA 60

Qy          61 FSIYILNLAAADFLFLSGRLIYSLLSFISIPHTISKILYPVMMFSYFAGLSFLSAVSTER 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 FSIYILNLAAADFLFLSGRLIYSLLSFISIPHTISKILYPVMMFSYFAGLSFLSAVSTER 120

Qy         121 CLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQTSDFIT 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 CLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQTSDFIT 180

Qy         181 VAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGIQFFLFL 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 VAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGIQFFLFL 240

Qy         241 WIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASE 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 WIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASE 300

Qy         301 VDEGGGQLPEEILELSGSRLEQ 322
               ||||||||||||||||||||||
Db         301 VDEGGGQLPEEILELSGSRLEQ 322


RESULT 3
US-09-995-225-20
; Sequence 20, Application US/09995225
; Publication No. US20020193584A1
; GENERAL INFORMATION:
;  APPLICANT: Chen, Ruoping
;  APPLICANT:  Chu, Zhi Liang
;  APPLICANT:  Dang, Huong T.
;  APPLICANT:  Lowitz, Kevin P.
;  APPLICANT:  Pride, Cameron
;  TITLE OF INVENTION: Endogenous And No. US20020193584A1-Endogenous Versions of Human G
;  TITLE OF INVENTION:  Receptors
;  FILE REFERENCE: AREN-0308
;  CURRENT APPLICATION NUMBER: US/09/995,225
;  CURRENT FILING DATE:  2001-11-26
;  PRIOR APPLICATION NUMBER: 09/170,496
;  PRIOR FILING DATE: 1998-10-13
;  PRIOR APPLICATION NUMBER: PCT/US99/23938
;  PRIOR FILING DATE: 1998-10-13
;  PRIOR APPLICATION NUMBER: 60/253,404
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/255,366
;  PRIOR FILING DATE: 2000-12-12
;  PRIOR APPLICATION NUMBER: 60/270,286
;  PRIOR FILING DATE: 2001-02-20
;  PRIOR APPLICATION NUMBER: 60/282,365
;  PRIOR FILING DATE: 2001-04-06
;  PRIOR APPLICATION NUMBER: 60/270,266
;  PRIOR FILING DATE: 2001-02-20
;  PRIOR APPLICATION NUMBER: 60/282,032
;  PRIOR FILING DATE: 2001-04-06
;  PRIOR APPLICATION NUMBER: 60/282,358
;  PRIOR FILING DATE: 2001-04-06
;  PRIOR APPLICATION NUMBER: 60/282,356
;  PRIOR FILING DATE: 2001-04-06
;  PRIOR APPLICATION NUMBER: 60/290,917
;  PRIOR FILING DATE: 2001-05-14
;  PRIOR APPLICATION NUMBER: 60/309,208
;  PRIOR FILING DATE: 2001-07-31
;  NUMBER OF SEQ ID NOS: 67
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
;   LENGTH: 322
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: No. US20020193584A1el Sequence
US-09-995-225-20

  Query Match           83.3%;  Score 1383;  DB 9;  Length 322;
  Best Local Similarity  83.5%;  Pred. No. 5.8e-117;
  Matches  269;  Conservative   16;  Mismatches   37;  Indels    0;  Gaps    0;

Qy           1 MDPTISTLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNA 60
               || ||  | |||||||| ||| ||||||| | |||||||| ||||||||||||||||||||
Db           1 MDSTIPVLGTELTPINGREETPCYKQTLSFTGLTCIVSLVALTGNAVVLWLLGCRMRRNA 60

Qy          61 FSIYILNLAAADFLFLSGRLIYSLLSFISIPHTISKILYPVMMFSYFAGLSFLSAVSTER 120
               ||||||| ||||||||| :| | |  |:| | ||||| ||| | || ||| |||:||||
Db          61 VSIYILNLVAADFLFLSGHIICSPLRLINIRHPISKILSPVMTFPYFIGLSMLSAISTER 120

Qy         121 CLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQTSDFIT 180
               |||:|||||| | || :||:|:|||||||||||||||| | |||||||| ||:||||||
Db         121 CLSILWPIWYHCRRPRYLSSVMCVLLWALSLLRSILEWMFCDFLFSGADSVWCETSDFIT 180

Qy         181 VAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGIQFFLFL 240
               :|||:|||||||||||||:|||||||:|||||||||||||||||||||||||||||: ||
Db         181 IAWLVFLCVVLCGSSLVLLVRILCGSRKMPLTRLYVTILLTVLVFLLCGLPFGIQWALFS 240

Qy         241 WIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASE 300
               ||:| :|||||||||||||||||||||||||||||||||||||||||||||||||| .|
Db         241 RIHLDWKVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDTPE 300

Qy         301 VDEGGGQLPEEILELSGSRLEQ 322
               |||||| ||:| ||||||||||
Db         301 VDEGGGWLPQETLELSGSRLEQ 322
```

RESULT 4
US-10-183-116-31
; Sequence 31, Application US/10183116
; Publication No. US20030092035A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David J.
; APPLICANT: Dong, Xinzhong
; APPLICANT: Zylka, Mark
; APPLICANT: Simon, Melvin
; APPLICANT: Han, Sang-kyou
; TITLE OF INVENTION: PAIN SIGNALING MOLECULES
; FILE REFERENCE: CALTE.4C1CP1
; CURRENT APPLICATION NUMBER: US/10/183,116
; CURRENT FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: US 60/222,344
; PRIOR FILING DATE: 2000-08-01
; PRIOR APPLICATION NUMBER: US 60/202,027
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 09/704,707
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: US 60/285,493
; PRIOR FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: US 09/849,869
; PRIOR FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 31
; LENGTH: 322
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-183-116-31

Query Match 83.3%; Score 1383; DB 9; Length 322;
Best Local Similarity 83.5%; Pred. No. 5.8e-117;
Matches 269; Conservative 16; Mismatches 37; Indels 0; Gaps 0;

```
Qy          1 MDPTISTLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNA 60
              || ||  | |||||||| ||| ||||||| | |||||||| ||||||||||||||||||||
Db          1 MDSTIPVLGTELTPINGREETPCYKQTLSFTGLTCIVSLVALTGNAVVLWLLGCRMRRNA 60

Qy         61 FSIYILNLAAADFLFLSGRLIYSLLSFISIPHTISKILYPVMMFSYFAGLSFLSAVSTER 120
              ||||||| ||||||||| :| | |  |:| | ||||| ||| | || ||| |||:||||
Db         61 VSIYILNLVAADFLFLSGHIICSPLRLINIRHPISKILSPVMTFPYFIGLSMLSAISTER 120

Qy        121 CLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQTSDFIT 180
              |||:|||||| | || :||:|:|||||||||||||||| | |||||||| ||:||||||
Db        121 CLSILWPIWYHCRRPRYLSSVMCVLLWALSLLRSILEWMFCDFLFSGADSVWCETSDFIT 180

Qy        181 VAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGIQFFLFL 240
              :|||:|||||||||||||:|||||||:|||||||||||||||||||||||||||||||: ||
Db        181 IAWLVFLCVVLCGSSLVLLVRILCGSRKMPLTRLYVTILLTVLVFLLCGLPFGIQWALFS 240

Qy        241 WIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASE 300
              ||:| :|||||||||||||||||||||||||||||||||||||||||||||||||||   |
Db        241 RIHLDWKVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDTPE 300

Qy        301 VDEGGGQLPEEILELSGSRLEQ 322
              |||||| ||:| ||||||||||
Db        301 VDEGGGWLPQETLELSGSRLEQ 322
```

RESULT 5
US-10-225-567A-674
; Sequence 674, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: LifeSpan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burmer, Glenna C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 674
; LENGTH: 322
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-225-567A-674

Query Match 83.3%; Score 1383; DB 9; Length 322;
Best Local Similarity 83.5%; Pred. No. 5.8e-117;
Matches 269; Conservative 16; Mismatches 37; Indels 0; Gaps 0;

```
Qy          1 MDPTISTLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNA 60
              || ||  | |||||||| ||| ||||||| | |||||||| ||||||||||||||||||||
Db          1 MDSTIPVLGTELTPINGREETPCYKQTLSFTGLTCIVSLVALTGNAVVLWLLGCRMRRNA 60

Qy         61 FSIYILNLAAADFLFLSGRLIYSLLSFISIPHTISKILYPVMMFSYFAGLSFLSAVSTER 120
              ||||||| ||||||||| :| | |  |:| | ||||| ||| | || ||| |||:||||
Db         61 VSIYILNLVAADFLFLSGHIICSPLRLINIRHPISKILSPVMTFPYFIGLSMLSAISTER 120

Qy        121 CLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQTSDFIT 180
              |||:|||||| | || :||:|:|||||||||||||||| | |||||||| ||:||||||
Db        121 CLSILWPIWYHCRRPRYLSSVMCVLLWALSLLRSILEWMFCDFLFSGADSVWCETSDFIT 180

Qy        181 VAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGIQFFLFL 240
              :|||:|||||||||||||:|||||||:|||||||||||||||||||||||||||||||: ||
Db        181 IAWLVFLCVVLCGSSLVLLVRILCGSRKMPLTRLYVTILLTVLVFLLCGLPFGIQWALFS 240

Qy        241 WIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASE 300
              ||:| :|||||||||||||||||||||||||||||||||||||||||||||||||||   |
Db        241 RIHLDWKVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDTPE 300

Qy        301 VDEGGGQLPEEILELSGSRLEQ 322
              |||||| ||:| ||||||||||
Db        301 VDEGGGWLPQETLELSGSRLEQ 322
```

RESULT 6
US-10-219-834-79
; Sequence 79, Application US/10219834
; Publication No. US20030096751A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: G-PROTEIN COUPLED RECEPTOR POLYNUCLEOTIDES AND METHODS OF USE THER
; FILE REFERENCE: D0191 NP
; CURRENT APPLICATION NUMBER: US/10/219,834
; CURRENT FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: US 60/313,658
; PRIOR FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: US 60/340,703
; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: US 60/318,675
; PRIOR FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: US 60/355,596
; PRIOR FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: US 60/333,417
; PRIOR FILING DATE: 2001-11-26
; PRIOR APPLICATION NUMBER: US 60/338,367
; PRIOR FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 79
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-219-834-79

Query Match 80.3%; Score 1334; DB 9; Length 314;
Best Local Similarity 81.1%; Pred. No. 1.5e-112;
Matches 261; Conservative 16; Mismatches 37; Indels 8; Gaps 1;

```
Qy          1 MDPTISTLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNA 60
              ||  ||   | ||||||||  ||| |||||||  | ||||||||  ||||||||||||||||||||
Db          1 MDSTIPVLGTELTPINGREETPCYKQTLSFTGLTCIVSLVALTGNAVVLWLLGCRMRRNA 60

Qy         61 FSIYILNLAAADFLFLSGRLIYSLLSFISIPHTISKILYPVMMFSYFAGLSFLSAVSTER 120
               |||||||  |||||||||  :| | |  |:| | |||||  ||| | ||  |||  |||:||||
Db         61 VSIYILNLVAADFLFLSGHIICSPLRLINIRHPISKILSPVMTFPYFIGLSMLSAISTER 120

Qy        121 CLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQTSDFIT 180
              |||:||||||  | ||  :||:|:|||||||||||||||  | ||||||||  ||:||||||
Db        121 CLSILWPIWYHCRRPRYLSSVMCVLLWALSLLRSILEWMFCDFLFSGADSVWCETSDFIT 180

Qy        181 VAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGIQFFLFL 240
              :|||:|||||||||||||:||||||||:|||||||||||||||||||||||||||||: ||
Db        181 IAWLVFLCVVLCGSSLVLLVRILCGSRKMPLTRLYVTILLTVLVFLLCGLPFGIQWALFS 240

Qy        241 WIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASE 300
              ||:| :|||||||||||||||||||||||||||||||||||||||||        |||  |
Db        241 RIHLDWKVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQN--------LQDTPE 292

Qy        301 VDEGGGQLPEEILELSGSRLEQ 322
              |||||| ||:| ||||||||||
Db        293 VDEGGGWLPQETLELSGSRLEQ 314
```

```
RESULT 7
US-09-995-225-18
; Sequence 18, Application US/09995225
; Publication No. US20020193584A1
; GENERAL INFORMATION:
;  APPLICANT: Chen, Ruoping
;  APPLICANT:  Chu, Zhi Liang
;  APPLICANT:  Dang, Huong T.
;  APPLICANT:  Lowitz, Kevin P.
;  APPLICANT:  Pride, Cameron
;  TITLE OF INVENTION: Endogenous And No. US20020193584A1-Endogenous Versions of Human G
;  TITLE OF INVENTION:  Receptors
;  FILE REFERENCE: AREN-0308
;  CURRENT APPLICATION NUMBER: US/09/995,225
;  CURRENT FILING DATE:  2001-11-26
;  PRIOR APPLICATION NUMBER: 09/170,496
;  PRIOR FILING DATE: 1998-10-13
;  PRIOR APPLICATION NUMBER: PCT/US99/23938
;  PRIOR FILING DATE: 1998-10-13
;  PRIOR APPLICATION NUMBER: 60/253,404
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/255,366
;  PRIOR FILING DATE: 2000-12-12
;  PRIOR APPLICATION NUMBER: 60/270,286
;  PRIOR FILING DATE: 2001-02-20
;  PRIOR APPLICATION NUMBER: 60/282,365
;  PRIOR FILING DATE: 2001-04-06
;  PRIOR APPLICATION NUMBER: 60/270,266
;  PRIOR FILING DATE: 2001-02-20
;  PRIOR APPLICATION NUMBER: 60/282,032
;  PRIOR FILING DATE: 2001-04-06
;  PRIOR APPLICATION NUMBER: 60/282,358
;  PRIOR FILING DATE: 2001-04-06
;  PRIOR APPLICATION NUMBER: 60/282,356
;  PRIOR FILING DATE: 2001-04-06
;  PRIOR APPLICATION NUMBER: 60/290,917
;  PRIOR FILING DATE: 2001-05-14
;  PRIOR APPLICATION NUMBER: 60/309,208
;  PRIOR FILING DATE: 2001-07-31
;  NUMBER OF SEQ ID NOS: 67
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
;   LENGTH: 322
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: No. US20020193584A1el Sequence
```

```
US-09-995-225-18

  Query Match             77.9%;  Score 1294;  DB 9;  Length 322;
  Best Local Similarity   79.7%;  Pred. No. 6.1e-109;
  Matches  255;  Conservative   20;  Mismatches   45;  Indels    0;  Gaps    0;

Qy          1 MDPTISTLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNA 60
              ||||:    |:|||||| ||| || |||| ||||||:|||||||||||||||| ||||||
Db          1 MDPTVPVFGTKLTPINGREETPCYNQTLSFTVLTCIISLVGLTGNAVVLWLLGYRMRRNA 60

Qy         61 FSIYILNLAAADFLFLSGRLIYSLLSFISIPHTISKILYPVMMFSYFAGLSFLSAVSTER 120
              |||||||||||||||| ::| | |  |:| | | |||  || | || ||| |||:||||
Db         61 VSIYILNLAAADFLFLSFQIIRSPLRLINISHLIRKILVSVMTFPYFTGLSMLSAISTER 120

Qy        121 CLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQTSDFIT 180
              |||||||||||| |||||||||||||| |||| |:|||  | ||||||||:||:|||||
Db        121 CLSVLWPLWYRCRRPTHLSAVVCVLLWGLSLLFSMLEWRFCDFLFSGADSSWCETSDFIP 180

Qy        181 VAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGIQFFLFL 240
              |||||||||||| ||||||:||||||||:||||||||||||||||||||||||    |
Db        181 VAWLIFLCVVLCVSSLVLLVRILCGSRKMPLTRLYVTILLTVLVFLLCGLPFGILGALIY 240

Qy        241 WIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASE 300
              :|:: |||:|||:|| : ||:||||||||||||||||||||||||||||||||||||  |
Db        241 RMHLNLEVLYCHVYLVCMSLSSLNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDKPE 300

Qy        301 VDEGGGQLPEEILELSGSRL 320
              ||:| |||||| ||||||||
Db        301 VDKGEGQLPEESLELSGSRL 320
```

```
RESULT 8
US-10-183-116-33
; Sequence 33, Application US/10183116
; Publication No. US20030092035A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David J.
;  APPLICANT:  Dong, Xinzhong
;  APPLICANT:  Zylka, Mark
;  APPLICANT:  Simon, Melvin
;  APPLICANT:  Han, Sang-kyou
;  TITLE OF INVENTION: PAIN SIGNALING MOLECULES
;  FILE REFERENCE: CALTE.4C1CP1
;  CURRENT APPLICATION NUMBER: US/10/183,116
;  CURRENT FILING DATE:  2002-06-26
;  PRIOR APPLICATION NUMBER: US 60/222,344
;  PRIOR FILING DATE: 2000-08-01
;  PRIOR APPLICATION NUMBER: US 60/202,027
;  PRIOR FILING DATE: 2000-05-04
;  PRIOR APPLICATION NUMBER: US 09/704,707
;  PRIOR FILING DATE: 2000-11-03
;  PRIOR APPLICATION NUMBER: US 60/285,493
;  PRIOR FILING DATE: 2001-04-19
;  PRIOR APPLICATION NUMBER: US 09/849,869
;  PRIOR FILING DATE: 2001-05-04
;  NUMBER OF SEQ ID NOS: 109
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 33
;   LENGTH: 322
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-183-116-33

  Query Match             77.9%;  Score 1294;  DB 9;  Length 322;
  Best Local Similarity   79.7%;  Pred. No. 6.1e-109;
  Matches  255;  Conservative   20;  Mismatches   45;  Indels    0;  Gaps    0;

Qy          1 MDPTISTLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNA 60
              ||||:    |:|||||| ||| || |||| ||||||:|||||||||||||||| ||||||
Db          1 MDPTVPVFGTKLTPINGREETPCYNQTLSFTVLTCIISLVGLTGNAVVLWLLGYRMRRNA 60

Qy         61 FSIYILNLAAADFLFLSGRLIYSLLSFISIPHTISKILYPVMMFSYFAGLSFLSAVSTER 120
              |||||||||||||||| ::| | |  |:| | | |||  || | || ||| |||:||||
```

```
Db         61 VSIYILNLAAADFLFLSFQIIRSPLRLINISHLIRKILVSVMTFPYFTGLSMLSAISTER 120

Qy        121 CLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQTSDFIT 180
              |||||||||||| |||||||||||||| |||| |:|||  | ||||||||:||:|||||
Db        121 CLSVLWPIWYRCRRPTHLSAVVCVLLWGLSLLFSMLEWRFCDFLFSGADSSWCETSDFIP 180

Qy        181 VAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGIQFFLFL 240
              |||||||||||| ||||||:||||||||:||||||||||||||||||||||||  |
Db        181 VAWLIFLCVVLCVSSLVLLVRILCGSRKMPLTRLYVTILLTVLVFLLCGLPFGILGALIY 240

Qy        241 WIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASE 300
              :|:: |||:|||:|| : ||:||||||||||||||||||||||||||||||||  |
Db        241 RMHLNLEVLYCHVYLVCMSLSSLNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDKPE 300

Qy        301 VDEGGGQLPEEILELSGSRL 320
              ||:| |||||| ||||||||
Db        301 VDKGEGQLPEESLELSGSRL 320


RESULT 9
US-10-225-567A-689
; Sequence 689, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
;  APPLICANT: LifeSpan Biosciences
;  APPLICANT:  Brown, Joseph P.
;  APPLICANT:  Burmer, Glenna C.
;  APPLICANT:  Roush, Christine L.
;  TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
;  FILE REFERENCE: 1920-4-4
;  CURRENT APPLICATION NUMBER: US/10/225,567A
;  CURRENT FILING DATE:  2001-12-19
;  PRIOR APPLICATION NUMBER: 60/257,144
;  PRIOR FILING DATE: 2000-12-19
;  NUMBER OF SEQ ID NOS: 2292
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 689
;   LENGTH: 322
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-225-567A-689

  Query Match           77.9%;  Score 1294;  DB 9;  Length 322;
  Best Local Similarity  79.7%;  Pred. No. 6.1e-109;
  Matches  255;  Conservative   20;  Mismatches   45;  Indels    0;  Gaps    0;

Qy          1 MDPTISTLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNA 60
              ||||:    |:|||||| ||| || |||| ||||||:|||||||||||||||| ||||||
Db          1 MDPTVPVFGTKLTPINGREETPCYNQTLSFTVLTCIISLVGLTGNAVVLWLLGYRMRRNA 60

Qy         61 FSIYILNLAAADFLFLSGRLIYSLLSFISIPHTISKILYPVMMFSYFAGLSFLSAVSTER 120
              |||||||||||||||| ::| | |  |:| | | |||  || | || ||| |||:||||
Db         61 VSIYILNLAAADFLFLSFQIIRSPLRLINISHLIRKILVSVMTFPYFTGLSMLSAISTER 120

Qy        121 CLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQTSDFIT 180
              |||||||||||| |||||||||||||| |||| |:|||  | ||||||||:||:|||||
Db        121 CLSVLWPIWYRCRRPTHLSAVVCVLLWGLSLLFSMLEWRFCDFLFSGADSSWCETSDFIP 180

Qy        181 VAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGIQFFLFL 240
              |||||||||||| ||||||:||||||||:||||||||||||||||||||||||  |
Db        181 VAWLIFLCVVLCVSSLVLLVRILCGSRKMPLTRLYVTILLTVLVFLLCGLPFGILGALIY 240

Qy        241 WIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASE 300
              :|:: |||:|||:|| : ||:||||||||||||||||||||||||||||||||  |
Db        241 RMHLNLEVLYCHVYLVCMSLSSLNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDKPE 300

Qy        301 VDEGGGQLPEEILELSGSRL 320
              ||:| |||||| ||||||||
Db        301 VDKGEGQLPEESLELSGSRL 320


RESULT 10
US-10-219-834-20
; Sequence 20, Application US/10219834
; Publication No. US20030096751A1
; GENERAL INFORMATION:
;  APPLICANT: Bristol-Myers Squibb Company
;  TITLE OF INVENTION: G-PROTEIN COUPLED RECEPTOR POLYNUCLEOTIDES AND METHODS OF USE THER
;  FILE REFERENCE: D0191 NP
;  CURRENT APPLICATION NUMBER: US/10/219,834
;  CURRENT FILING DATE:  2002-08-15
;  PRIOR APPLICATION NUMBER: US 60/313,658
;  PRIOR FILING DATE: 2001-08-20
;  PRIOR APPLICATION NUMBER: US 60/340,703
;  PRIOR FILING DATE: 2001-10-30
;  PRIOR APPLICATION NUMBER: US 60/318,675
;  PRIOR FILING DATE: 2001-09-12
;  PRIOR APPLICATION NUMBER: US 60/355,596
;  PRIOR FILING DATE: 2002-02-06
;  PRIOR APPLICATION NUMBER: US 60/333,417
;  PRIOR FILING DATE: 2001-11-26
;  PRIOR APPLICATION NUMBER: US 60/338,367
;  PRIOR FILING DATE: 2001-12-06
;  NUMBER OF SEQ ID NOS: 192
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
;   LENGTH: 328
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-219-834-20

  Query Match           73.0%;  Score 1212;  DB 9;  Length 328;
  Best Local Similarity  74.7%;  Pred. No. 1.6e-101;
  Matches  245;  Conservative   17;  Mismatches   60;  Indels    6;  Gaps    2;

Qy          1 MDPTISTLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNA 60
              || ||  | |||||||| ||| ||||||| | |||||||| ||||||||||||||||||||
Db          1 MDSTIPVLGTELTPINGREETPCYKQTLSFTGLTCIVSLVALTGNAVVLWLLGCRMRRNA 60

Qy         61 FSIYILNLAAADFLFLSGRLIYSLL----SFISIPHTISKILYPVMMFSYFAGLSFLSAV 116
              ||||||| |||||||  ::|  |:     | ||          ||  :| |||| || |
Db         61 VSIYILNLVAADFLFLCFQIINCLVYLSNFFCSISINFPSFFTTVMTCAYLAGLSMLSTV 120

Qy        117 STERCLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQTS 176
              |||||||||||||||| || |||||||||||||||| ||||   ||||||  || ||||
Db        121 STERCLSVLWPIWYRCRRPRHLSAVVCVLLWALSLLLSILEGKFCGFLFSDGDSGWCQTF 180

Qy        177 DFITVAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGIQF 236
              |||| |||||_:|||||||| ||:||||||| :|||||||:||||||||||||||||||:
Db        181 DFITAAWLIFLFMVLCGSSLALLVRILCGSRGLPLTRLYLTILLTVLVFLLCGLPFGIQW 240

Qy        237 FLFLWIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQN--LKLVLQRA 294
              || |||  | :|||||:| ||: ||:|||||||||||||||::   |   ||| ||||
Db        241 FLILWIWKDSDVLFCHIHPVSVVLSSLNSSANPIIYFFVGSFRKQWRLQQPILKLALQRA 300

Qy        295 LQDASEVDEGGGQLPEEILELSGSRLEQ 322
              ||| :||||||| ||:| ||||||||||
Db        301 LQDIAEVDEGGGWLPQETLELSGSRLEQ 328


RESULT 11
US-10-183-116-18
; Sequence 18, Application US/10183116
; Publication No. US20030092035A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David J.
;  APPLICANT:  Dong, Xinzhong
;  APPLICANT:  Zylka, Mark
;  APPLICANT:  Simon, Melvin
;  APPLICANT:  Han, Sang-kyou
;  TITLE OF INVENTION: PAIN SIGNALING MOLECULES
;  FILE REFERENCE: CALTE.4C1CP1
;  CURRENT APPLICATION NUMBER: US/10/183,116
;  CURRENT FILING DATE:  2002-06-26
;  PRIOR APPLICATION NUMBER: US 60/222,344
;  PRIOR FILING DATE: 2000-08-01
```

```
;  PRIOR APPLICATION NUMBER: US 60/202,027
;  PRIOR FILING DATE: 2000-05-04
;  PRIOR APPLICATION NUMBER: US 09/704,707
;  PRIOR FILING DATE: 2000-11-03
;  PRIOR APPLICATION NUMBER: US 60/285,493
;  PRIOR FILING DATE: 2001-04-19
;  PRIOR APPLICATION NUMBER: US 09/849,869
;  PRIOR FILING DATE: 2001-05-04
;  NUMBER OF SEQ ID NOS: 109
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 18
;   LENGTH: 330
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-183-116-18

  Query Match             62.5%;  Score 1038.5;  DB 9;  Length 330;
  Best Local Similarity   64.7%;  Pred. No. 7e-86;
  Matches  213;  Conservative   27;  Mismatches   80;  Indels    9;  Gaps    3;

Qy          1 MDPTISTLDTELTPINGTEET---LCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMR 57
              ||||      || | :|| ::    || |:||    |   ::|||| ||  |||||| |||
Db          1 MDPTTPAWGTESTTVNGNDQALLLLCGKETLIPVFLILFIALVGLVGNGFVLWLLGFRMR 60

Qy         58 RNAFSIYILNLAAADFLFLSGRLIYSLL----SFISIPHTISKILYPVMMFSYFAGLSFL 113
              |||||:|:|:|| ||||||  ::|  |:     | ||           ||  :| |||| |
Db         61 RNAFSVYVLSLAGADFLFLCFQIINCLVYLSNFFCSISINFPSFFTTVMTCAYLAGLSML 120

Qy        114 SAVSTERCLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWC 173
              | |||||||||||||||| || |||||||||||||||| ||||   ||||||  || ||
Db        121 STVSTERCLSVLWPIWYRCRRPRHLSAVVCVLLWALSLLLSILEGKFCGFLFSDGDSGWC 180

Qy        174 QTSDFITVAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFG 233
              || |||| |||||| :||||||| ||:||||||| :||||||:|||||||||||||||||
Db        181 QTFDFITAAWLIFLFMVLCGSSLALLVRILCGSRGLPLTRLYLTILLTVLVFLLCGLPFG 240

Qy        234 IQFFLFLWIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQN--LKLVL 291
              ||:|| |||  | :|||||:| ||: ||:||||||||||||||::   |   ||| |
Db        241 IQWFLILWIWKDSDVLFCHIHPVSVVLSSLNSSANPIIYFFVGSFRKQWRLQQPILKLAL 300

Qy        292 QRALQDASEVDEGGGQLPEEILELSGSRL 320
              |||||| :|||   |   :   |:| | |
Db        301 QRALQDIAEVDHSEGCFRQGTPEMSRSSL 329


RESULT 12
US-10-079-384-12
; Sequence 12, Application US/10079384
; Publication No. US20030108986A1
; GENERAL INFORMATION:
;  APPLICANT: Communi, Didier
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS COMPRISING G-PROTEIN COUPLED RECEPTORS
;  FILE REFERENCE: 9409/2132
;  CURRENT APPLICATION NUMBER: US/10/079,384
;  CURRENT FILING DATE:  2002-02-20
;  PRIOR APPLICATION NUMBER: US 09/885,453
;  PRIOR FILING DATE: 2001-06-20
;  NUMBER OF SEQ ID NOS: 50
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
;   LENGTH: 330
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-079-384-12

  Query Match             62.5%;  Score 1038.5;  DB 9;  Length 330;
  Best Local Similarity   64.7%;  Pred. No. 7e-86;
  Matches  213;  Conservative   27;  Mismatches   80;  Indels    9;  Gaps    3;

Qy          1 MDPTISTLDTELTPINGTEET---LCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMR 57
              ||||      || | :|| ::    || |:||    |   ::|||| ||  |||||| |||
Db          1 MDPTTPAWGTESTTVNGNDQALLLLCGKETLIPVFLILFIALVGLVGNGFVLWLLGFRMR 60

Qy         58 RNAFSIYILNLAAADFLFLSGRLIYSLL----SFISIPHTISKILYPVMMFSYFAGLSFL 113
              |||||:|:|:|| ||||||  ::|  |:     | ||           ||  :| |||| |
Db         61 RNAFSVYVLSLAGADFLFLCFQIINCLVYLSNFFCSISINFPSFFTTVMTCAYLAGLSML 120

Qy        114 SAVSTERCLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWC 173
              | |||||||||||||||| || |||||||||||||||| ||||   ||||||  || ||
Db        121 STVSTERCLSVLWPIWYRCRRPRHLSAVVCVLLWALSLLLSILEGKFCGFLFSDGDSGWC 180

Qy        174 QTSDFITVAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFG 233
              || |||| |||||| :||||||| ||:||||||| :||||||:|||||||||||||||||
Db        181 QTFDFITAAWLIFLFMVLCGSSLALLVRILCGSRGLPLTRLYLTILLTVLVFLLCGLPFG 240

Qy        234 IQFFLFLWIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQN--LKLVL 291
              ||:|| |||  | :|||||:| ||: ||:||||||||||||||::   |   ||| |
Db        241 IQWFLILWIWKDSDVLFCHIHPVSVVLSSLNSSANPIIYFFVGSFRKQWRLQQPILKLAL 300

Qy        292 QRALQDASEVDEGGGQLPEEILELSGSRL 320
              |||||| :|||   |   :   |:| | |
Db        301 QRALQDIAEVDHSEGCFRQGTPEMSRSSL 329


RESULT 13
US-10-225-567A-649
; Sequence 649, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
;  APPLICANT: LifeSpan Biosciences
;  APPLICANT:  Brown, Joseph P.
;  APPLICANT:  Burmer, Glenna C.
;  APPLICANT:  Roush, Christine L.
;  TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
;  FILE REFERENCE: 1920-4-4
;  CURRENT APPLICATION NUMBER: US/10/225,567A
;  CURRENT FILING DATE:  2001-12-19
;  PRIOR APPLICATION NUMBER: 60/257,144
;  PRIOR FILING DATE: 2000-12-19
;  NUMBER OF SEQ ID NOS: 2292
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 649
;   LENGTH: 330
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-225-567A-649

  Query Match             62.5%;  Score 1038.5;  DB 9;  Length 330;
  Best Local Similarity   64.7%;  Pred. No. 7e-86;
  Matches  213;  Conservative   27;  Mismatches   80;  Indels    9;  Gaps    3;

Qy          1 MDPTISTLDTELTPINGTEET---LCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMR 57
              ||||      || | :|| ::    || |:||    |   ::|||| ||  |||||| |||
Db          1 MDPTTPAWGTESTTVNGNDQALLLLCGKETLIPVFLILFIALVGLVGNGFVLWLLGFRMR 60

Qy         58 RNAFSIYILNLAAADFLFLSGRLIYSLL----SFISIPHTISKILYPVMMFSYFAGLSFL 113
              |||||:|:|:|| ||||||  ::|  |:     | ||           ||  :| |||| |
Db         61 RNAFSVYVLSLAGADFLFLCFQIINCLVYLSNFFCSISINFPSFFTTVMTCAYLAGLSML 120

Qy        114 SAVSTERCLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWC 173
              | |||||||||||||||| || |||||||||||||||| ||||   ||||||  || ||
Db        121 STVSTERCLSVLWPIWYRCRRPRHLSAVVCVLLWALSLLLSILEGKFCGFLFSDGDSGWC 180

Qy        174 QTSDFITVAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFG 233
              || |||| |||||| :||||||| ||:||||||| :||||||:|||||||||||||||||
Db        181 QTFDFITAAWLIFLFMVLCGSSLALLVRILCGSRGLPLTRLYLTILLTVLVFLLCGLPFG 240

Qy        234 IQFFLFLWIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQN--LKLVL 291
              ||:|| |||  | :|||||:| ||: ||:||||||||||||||::   |   ||| |
Db        241 IQWFLILWIWKDSDVLFCHIHPVSVVLSSLNSSANPIIYFFVGSFRKQWRLQQPILKLAL 300

Qy        292 QRALQDASEVDEGGGQLPEEILELSGSRL 320
              |||||| :|||   |   :   |:| | |
Db        301 QRALQDIAEVDHSEGCFRQGTPEMSRSSL 329
```

```
RESULT 14
US-09-826-508-12
; Sequence 12, Application US/09826508
; Patent No. US20010025099A1
; GENERAL INFORMATION:
;  APPLICANT: Nabil Elshourbagy
;  APPLICANT:  Lisa Vawter
;  TITLE OF INVENTION: G Protein-Coupled Receptor Polypeptides
;  TITLE OF INVENTION:  and Polynucleotides
;  FILE REFERENCE: GP-70744USB
;  CURRENT APPLICATION NUMBER: US/09/826,508
;  CURRENT FILING DATE:  2001-04-05
;  NUMBER OF SEQ ID NOS: 40
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 12
;   LENGTH: 330
;   TYPE: PRT
;   ORGANISM: HOMO SAPIENS
US-09-826-508-12

  Query Match             62.5%;  Score 1038.5;  DB 10;  Length 330;
  Best Local Similarity   64.7%;  Pred. No. 7e-86;
  Matches  213;  Conservative   27;  Mismatches   80;  Indels    9;  Gaps    3;

Qy          1 MDPTISTLDTELTPINGTEET---LCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMR 57
              |||      || | :|| ::    || |:||     |   ::|||| ||  |||||| |||
Db          1 MDPTTPAWGTESTTVNGNDQALLLLCGKETLIPVFLILFIALVGLVGNGFVLWLLGFRMR 60

Qy         58 RNAFSIYILNLAAADFLFLSGRLIYSLL----SFISIPHTISKILYPVMMFSYFAGLSFL 113
              |||||:|:|:|| ||||||  ::|  |:     | ||          ||  :| |||| |
Db         61 RNAFSVYVLSLAGADFLFLCFQIINCLVYLSNFFCSISINFPSFFTTVMTCAYLAGLSML 120

Qy        114 SAVSTERCLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWC 173
              | |||||||||||||||| || |||||||||||||||| ||||   ||||||  || ||
Db        121 STVSTERCLSVLWPIWYRCRRPRHLSAVVCVLLWALSLLLSILEGKFCGFLFSDGDSGWC 180

Qy        174 QTSDFITVAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFG 233
              || |||| |||||| :||||||| ||:||||||| :||||||:|||||||||||||||||
Db        181 QTFDFITAAWLIFLFMVLCGSSLALLVRILCGSRGLPLTRLYLTILLTVLVFLLCGLPFG 240

Qy        234 IQFFLFLWIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQN--LKLVL 291
              ||:|| |||   | :|||||:| ||: ||:||||||||||||||::   |   ||| |
Db        241 IQWFLILWIWKDSDVLFCHIHPVSVVLSSLNSSANPIIYFFVGSFRKQWRLQQPILKLAL 300

Qy        292 QRALQDASEVDEGGGQLPEEILELSGSRL 320
              |||||| :|||   |   :   |:| | |
Db        301 QRALQDIAEVDHSEGCFRQGTPEMSRSSL 329


RESULT 15
US-09-750-373-24
; Sequence 24, Application US/09750373
; Patent No. US20020062013A1
; GENERAL INFORMATION:
;  APPLICANT: Lind, Peter
;  APPLICANT:  Wood, Linda S.
;  APPLICANT:  Hiebsch, Ronald
;  APPLICANT:  Ruff, Valerie
;  APPLICANT:  Lindberg, Eleni
;  APPLICANT:  Parodi, Luis A.
;  APPLICANT:  Vogeli, Gabriel
;  TITLE OF INVENTION: No. US20020062013A1el G Protein Coupled Receptors
;  FILE REFERENCE: PHRM-0300
;  CURRENT APPLICATION NUMBER: US/09/750,373
;  CURRENT FILING DATE:  2000-12-28
;  PRIOR APPLICATION NUMBER: 60/184,305
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: 60/188,880
;  PRIOR FILING DATE: 2000-03-13
;  PRIOR APPLICATION NUMBER: 60/219,492
;  PRIOR FILING DATE: 2000-07-20
;  PRIOR APPLICATION NUMBER: 60/173,339
;  PRIOR FILING DATE: 1999-12-28
;  PRIOR APPLICATION NUMBER: 60/224,321
;  PRIOR FILING DATE: 2000-08-11
;  PRIOR APPLICATION NUMBER: 60/200,534
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: 60/239,062
;  PRIOR FILING DATE: 2000-10-09
;  NUMBER OF SEQ ID NOS: 56
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 24
;   LENGTH: 330
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-750-373-24

  Query Match             62.5%;  Score 1038.5;  DB 10;  Length 330;
  Best Local Similarity   64.7%;  Pred. No. 7e-86;
  Matches  213;  Conservative   27;  Mismatches   80;  Indels    9;  Gaps    3;

Qy          1 MDPTISTLDTELTPINGTEET---LCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMR 57
              |||      || | :|| ::    || |:||     |   ::|||| ||  |||||| |||
Db          1 MDPTTPAWGTESTTVNGNDQALLLLCGKETLIPVFLILFIALVGLVGNGFVLWLLGFRMR 60

Qy         58 RNAFSIYILNLAAADFLFLSGRLIYSLL----SFISIPHTISKILYPVMMFSYFAGLSFL 113
              |||||:|:|:|| ||||||  ::|  |:     | ||          ||  :| |||| |
Db         61 RNAFSVYVLSLAGADFLFLCFQIINCLVYLSNFFCSISINFPSFFTTVMTCAYLAGLSML 120

Qy        114 SAVSTERCLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWC 173
              | |||||||||||||||| || |||||||||||||||| ||||   ||||||  || ||
Db        121 STVSTERCLSVLWPIWYRCRRPRHLSAVVCVLLWALSLLLSILEGKFCGFLFSDGDSGWC 180

Qy        174 QTSDFITVAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFG 233
              || |||| |||||| :||||||| ||:||||||| :||||||:|||||||||||||||||
Db        181 QTFDFITAAWLIFLFMVLCGSSLALLVRILCGSRGLPLTRLYLTILLTVLVFLLCGLPFG 240

Qy        234 IQFFLFLWIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQN--LKLVL 291
              ||:|| |||   | :|||||:| ||: ||:||||||||||||||::   |   ||| |
Db        241 IQWFLILWIWKDSDVLFCHIHPVSVVLSSLNSSANPIIYFFVGSFRKQWRLQQPILKLAL 300

Qy        292 QRALQDASEVDEGGGQLPEEILELSGSRL 320
              |||||| :|||   |   :   |:| | |
Db        301 QRALQDIAEVDHSEGCFRQGTPEMSRSSL 329


Search completed: July  2, 2003, 19:23:45
Job time : 25 secs
```

GenCore version 5.1.6


OM protein - protein search, using sw model

Run on: July 2, 2003, 19:19:30 ; Search time 14 Seconds
(without alignments)
676.727 Million cell updates/sec

Title: US-09-849-869A-16
Perfect score: 1661
Sequence: 1 MDPTISTLDTELTPINGTEE.....EGGGQLPEEILELSGSRLEQ 322

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 451.5 | 27.2 | 325 | 6 | 5320941-2 | Patent No. 5320941 |
| 2 | 410.5 | 24.7 | 282 | 1 | US-08-118-270-52 | Sequence 52, Appl |
| 3 | 410.5 | 24.7 | 282 | 5 | PCT-US93-08528-52 | Sequence 52, Appl |
| 4 | 384 | 23.1 | 298 | 1 | US-08-118-270-76 | Sequence 76, Appl |
| 5 | 384 | 23.1 | 298 | 5 | PCT-US93-08528-76 | Sequence 76, Appl |
| 6 | 279.5 | 16.8 | 395 | 3 | US-08-981-825-6 | Sequence 6, Appli |
| 7 | 279.5 | 16.8 | 395 | 4 | US-09-480-784-6 | Sequence 6, Appli |
| 8 | 250 | 15.1 | 350 | 2 | US-08-458-970A-9 | Sequence 9, Appli |
| 9 | 243.5 | 14.7 | 369 | 2 | US-08-411-859-3 | Sequence 3, Appli |
| 10 | 243.5 | 14.7 | 369 | 4 | US-08-387-707-9 | Sequence 9, Appli |
| 11 | 243.5 | 14.7 | 369 | 4 | US-08-405-271A-9 | Sequence 9, Appli |
| 12 | 241.5 | 14.5 | 369 | 1 | US-07-816-283-8 | Sequence 8, Appli |
| 13 | 241.5 | 14.5 | 369 | 1 | US-08-417-103-8 | Sequence 8, Appli |
| 14 | 241.5 | 14.5 | 369 | 4 | US-08-120-601B-9 | Sequence 9, Appli |
| 15 | 236 | 14.2 | 364 | 2 | US-08-458-970A-10 | Sequence 10, Appl |
| 16 | 226 | 13.6 | 354 | 1 | US-07-759-568-2 | Sequence 2, Appli |
| 17 | 224 | 13.5 | 381 | 1 | US-08-467-125-2 | Sequence 2, Appli |
| 18 | 224 | 13.5 | 381 | 2 | US-08-911-320A-2 | Sequence 2, Appli |
| 19 | 224 | 13.5 | 381 | 4 | US-09-217-101-2 | Sequence 2, Appli |
| 20 | 222 | 13.4 | 380 | 3 | US-08-676-351-5 | Sequence 5, Appli |
| 21 | 219.5 | 13.2 | 369 | 1 | US-07-816-283-6 | Sequence 6, Appli |
| 22 | 219.5 | 13.2 | 369 | 1 | US-08-417-103-6 | Sequence 6, Appli |
| 23 | 219.5 | 13.2 | 369 | 1 | US-08-417-103-16 | Sequence 16, Appl |
| 24 | 219.5 | 13.2 | 384 | 3 | US-09-071-434-3 | Sequence 3, Appli |
| 25 | 219 | 13.2 | 304 | 1 | US-08-118-270-35 | Sequence 35, Appl |
| 26 | 219 | 13.2 | 304 | 5 | PCT-US93-08528-35 | Sequence 35, Appl |
| 27 | 218.5 | 13.2 | 333 | 1 | US-08-148-215A-4 | Sequence 4, Appli |
| 28 | 218.5 | 13.2 | 391 | 4 | US-08-120-601B-8 | Sequence 8, Appli |
| 29 | 218 | 13.1 | 380 | 1 | US-08-149-093A-7 | Sequence 7, Appli |
| 30 | 218 | 13.1 | 380 | 2 | US-08-911-245-7 | Sequence 7, Appli |
| 31 | 218 | 13.1 | 380 | 4 | US-08-188-275A-5 | Sequence 5, Appli |
| 32 | 218 | 13.1 | 380 | 4 | US-09-510-473-7 | Sequence 7, Appli |
| 33 | 218 | 13.1 | 380 | 4 | US-09-351-198-5 | Sequence 5, Appli |
| 34 | 218 | 13.1 | 380 | 4 | US-09-113-426-5 | Sequence 5, Appli |
| 35 | 217.5 | 13.1 | 350 | 1 | US-07-759-568-3 | Sequence 3, Appli |
| 36 | 217 | 13.1 | 259 | 4 | US-09-261-599B-3 | Sequence 3, Appli |
| 37 | 217 | 13.1 | 259 | 4 | US-09-456-455A-3 | Sequence 3, Appli |
| 38 | 217 | 13.1 | 380 | 3 | US-08-147-592A-2 | Sequence 2, Appli |
| 39 | 217 | 13.1 | 380 | 4 | US-08-292-694A-2 | Sequence 2, Appli |
| 40 | 215 | 12.9 | 428 | 1 | US-07-816-283-12 | Sequence 12, Appl |
| 41 | 215 | 12.9 | 428 | 1 | US-08-417-103-12 | Sequence 12, Appl |
| 42 | 214 | 12.9 | 394 | 4 | US-08-405-271A-22 | Sequence 22, Appl |
| 43 | 213 | 12.8 | 370 | 3 | US-08-781-250-2 | Sequence 2, Appli |
| 44 | 212.5 | 12.8 | 315 | 1 | US-08-118-270-34 | Sequence 34, Appl |
| 45 | 212.5 | 12.8 | 315 | 5 | PCT-US93-08528-34 | Sequence 34, Appl |

ALIGNMENTS

RESULT 1
5320941-2
;Patent No. 5320941
; APPLICANT: Young, Dallan;Wigler, Michael H.;Fasano
;Ottavio
; TITLE OF INVENTION: DNA SEQUENCES ENCODING MAS ONHCOGENE,
;POLYPETIDES ENCODED THEREFROM AND DIAGNOSTIC AND OTHER METHODS
;BASED THEREFROM
; NUMBER OF SEQUENCES: 2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/06/872,087
; FILING DATE: 06-JUN-1986
;SEQ ID NO:2:
; LENGTH: 325
5320941-2

Query Match 27.2%; Score 451.5; DB 6; Length 325;
Best Local Similarity 38.4%; Pred. No. 3.4e-29;
Matches 103; Conservative 59; Mismatches 95; Indels 11; Gaps 4;

```
Qy         37 VSLVGLTGNAVVLWLLGCRMRRNAFSIYILNLAAADFLFLSGRLI----YSLLSFISIPH 92
              :| ||   | ::|| |  ||||| |::|| :|: ||   |    |    |:|   :|  |
Db         42 ISPVGFVENGILLWFLCFRMRRNPFTVYITHLSIADISLLFCIFILSIDYALDYELSSGH 101

Qy         93 TISKILYPV-MMFSYFAGLSFLSAVSTERCLSVLWPIWYRCHRPTHLSAVVCVLLWALSL 151
                 : :   |  :| |  ||  |:|:| |||||||:||||||||| : ||:|| ||||||
Db        102 YYTIVTLSVTFLFGYNTGLYLLTAISVERCLSVLYPIWYRCHRPKYQSALVCALLWALSC 161

Qy        152 LRSILEWMLCGFLFSGADSAWCQTSDFITVAWLIFLCV--VLCGSSLVLLIRILCGSRKI 209
              | : :|:::|       :||     :  | :| | ||    ::  || :|:::|   :
Db        162 LVTTMEYVMCIHTEEESDSPNDCPAVIIFIAILSFLVFTGLMLVSSTILVVKIPKNTWAS 221

Qy        210 PLTRLYVTILLTVLVFLLCGLPFGIQFFLFLWIHVDREVLFCHVHLVSIFLSALNSSANP 269
                ::||: |::|:::||:   | ::  |   :: :    | ::| :|:  | :|||||
Db        222 HSSKLYIVIMVTIIIFLI----FAMRMRLLYLLYYEYWSTFGNLHDISLLFSTINSSANN 277

Qy        270 IIYFFVGSFRQRQNRQNLKLVLQRALQD 297
              ||||||| :::: :|:||:|  || :|
Db        278 FIYFFVGSSKKKRFKQSLKVVTTRAFKD 305
```

RESULT 2
US-08-118-270-52
; Sequence 52, Application US/08118270
; Patent No. 5508384
; GENERAL INFORMATION:
; APPLICANT: Murphy, Randall B.
; APPLICANT: Schuster, David I.
; TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
; TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF

```
;    NUMBER OF SEQUENCES:  348
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/118,270
;      FILING DATE:  09-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/943,236
;      FILING DATE:  10-SEP-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Townsend, Kevin G.
;      REGISTRATION NUMBER:  34,033
;      REFERENCE/DOCKET NUMBER:  MURPHY=2A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  52:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  282 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-118-270-52

  Query Match             24.7%;  Score 410.5;  DB 1;  Length 282;
  Best Local Similarity   37.2%;  Pred. No. 5.8e-26;
  Matches  100;  Conservative   58;  Mismatches   88;  Indels   23;  Gaps    9;

Qy         37 VSLVGLTGNAVVLWLLGCRMRRNAFSIYI-LNLAAADFLF----LSGRLIYSLLSFISIP 91
              :| ||   | ::|| | |      |::|  |::|    ||    ||  : |:|   :|
Db          9 ISPVGFVENGILLWFL-C-----FFTVYTHLSIADISLLFCIFILS--IDYALDYELSSG 60

Qy         92 HTISKILYPV-MMFSYFAGLSFLSAVSTERCLSVLWPIWYRCHRPTHLSAVVCVLLWALS 150
              |  : :   |  :| |  ||  |:|:| |||||||:||||||||| : ||:|| ||||||
Db         61 HYYTIVTLSVTFLFGYNTGLYLLTAISVERCLSVLYPIWYRCHRPKYQSALVCALLWALS 120

Qy        151 LLRSILEWMLCGFLFSGADSAWCQTSDFITVAWLIFLCVVLCGSSLVLLIRILCGSRKIP 210
              | : : :::|   |  : |     :  | :| | ||      || :|:::|   :
Db        121 CLVTTM-YVMCIDRFEESHSRNDCRAVIIFIAILSFLVFTPSVSSTILVVKIRKNTWASH 179

Qy        211 LTRLYVTILLTVLVFLLCGLPFGIQFFLF--LWIHVDREVLFCHVHLVSIFLSALNSSAN 268
              ::||: |::|:::||:  :|  : : |:   |       | ::| :|:  | :|||||
Db        180 SSKLYIVIMVTIIIFLIFAMPMRLLYLLYYEYW------STFGNLHHISLLFSTINSSAN 233

Qy        269 PIIYFFVGSFRQRQNRQNLKLVLQRALQD 297
              | ||||||| :::: :::||:|| || :|
Db        234 PFIYFFVGSSKKKRFKESLKVVLTRAFKD 262


RESULT 3
PCT-US93-08528-52
; Sequence 52, Application PC/TUS9308528
;  GENERAL INFORMATION:
;    APPLICANT:  New York University
;    TITLE OF INVENTION:  POLYPEPTIDES OF G-COUPLED PROTEIN
;    TITLE OF INVENTION:  RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
;    NUMBER OF SEQUENCES:  348
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/08528
;      FILING DATE:  09-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/943,236
;      FILING DATE:  10-SEP-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Townsend, Kevin G.
;      REGISTRATION NUMBER:  34,033
;      REFERENCE/DOCKET NUMBER:  MURPHY=2 PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  52:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  282 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
PCT-US93-08528-52

  Query Match             24.7%;  Score 410.5;  DB 5;  Length 282;
  Best Local Similarity   37.2%;  Pred. No. 5.8e-26;
  Matches  100;  Conservative   58;  Mismatches   88;  Indels   23;  Gaps    9;

Qy         37 VSLVGLTGNAVVLWLLGCRMRRNAFSIYI-LNLAAADFLF----LSGRLIYSLLSFISIP 91
              :| ||   | ::|| | |      |::|  |::|    ||    ||  : |:|   :|
Db          9 ISPVGFVENGILLWFL-C-----FFTVYTHLSIADISLLFCIFILS--IDYALDYELSSG 60

Qy         92 HTISKILYPV-MMFSYFAGLSFLSAVSTERCLSVLWPIWYRCHRPTHLSAVVCVLLWALS 150
              |  : :   |  :| |  ||  |:|:| |||||||:||||||||| : ||:|| ||||||
Db         61 HYYTIVTLSVTFLFGYNTGLYLLTAISVERCLSVLYPIWYRCHRPKYQSALVCALLWALS 120

Qy        151 LLRSILEWMLCGFLFSGADSAWCQTSDFITVAWLIFLCVVLCGSSLVLLIRILCGSRKIP 210
              | : : :::|   |  : |     :  | :| | ||      || :|:::|   :
Db        121 CLVTTM-YVMCIDRFEESHSRNDCRAVIIFIAILSFLVFTPSVSSTILVVKIRKNTWASH 179

Qy        211 LTRLYVTILLTVLVFLLCGLPFGIQFFLF--LWIHVDREVLFCHVHLVSIFLSALNSSAN 268
              ::||: |::|:::||:  :|  : : |:   |       | ::| :|:  | :|||||
Db        180 SSKLYIVIMVTIIIFLIFAMPMRLLYLLYYEYW------STFGNLHHISLLFSTINSSAN 233

Qy        269 PIIYFFVGSFRQRQNRQNLKLVLQRALQD 297
              | ||||||| :::: :::||:|| || :|
Db        234 PFIYFFVGSSKKKRFKESLKVVLTRAFKD 262


RESULT 4
US-08-118-270-76
; Sequence 76, Application US/08118270
; Patent No. 5508384
;  GENERAL INFORMATION:
;    APPLICANT:  Murphy, Randall B.
;    APPLICANT:  Schuster, David I.
;    TITLE OF INVENTION:  POLYPEPTIDES OF G-COUPLED PROTEIN
;    TITLE OF INVENTION:  RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
;    NUMBER OF SEQUENCES:  348
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
```

```
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/118,270
;      FILING DATE:  09-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/943,236
;      FILING DATE:  10-SEP-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Townsend, Kevin G.
;      REGISTRATION NUMBER:  34,033
;      REFERENCE/DOCKET NUMBER:  MURPHY=2A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  76:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  298 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-118-270-76

  Query Match             23.1%;  Score 384;  DB 1;  Length 298;
  Best Local Similarity   32.7%;  Pred. No. 8.4e-24;
  Matches   97;  Conservative   60;  Mismatches  102;  Indels   38;  Gaps   11;

Qy         36 IVSLVGLTGNAVVLWLLGCRMRRNAFSIYI--LNLAAADFLFLSGRLIYSLL---SFI-S 89
              :: | || || :||| |  ::|  ||||| |::|:|| ::|  : : :||   :|: |
Db          8 LLCLCGLVGNGLVLWFFGFSIKRTPFSIYIYFLHIASADGIYLFSKAVIALLNMGTFLGS 67

Qy         90 IPHTISKILYPVMMFSYFAGLSFLSAVSTERCLSVLWPIWYRCHRPTHLSAVVCVLLWAL 149
               |  : ::   :: ~::|||:| | |:| |||:||::|:||   ||  ||| || ||| |
Db         68 FPDYVRRV-SRIVGLTFFAGVSLLPAISIERCVSVIFPMWYWRRRPKRLSAGVCALLWLL 126

Qy        150 SLLRSILEWMLCGFLFSGADSAWCQTSDFITVAWLIF-----LCVVLCGSSLVLLIRILC 204
              | |.: :    |  |    |    |   :  |:|     : |: |    : |: : |
Db        127 SFLVTSIHNYFC-LLGHEASGTACLNMDISLLGILLFFLFCPIMVLPC----IALLHVEC 181

Qy        205 GSRKIPLT-RLYVTILLTVLVFLLCGLPFGIQFFLFLWI--------HVDREVLFCHVHL 255
               :|:   : :|   :|  | |||:  :  || :||| |:          |::  |
Db        182 RARRRQRSAKLNHVVLAIVSVFLVSSIYLGIDWFLF-WVFQIPAPFPEYVRDLCIC---- 236

Qy        256 VSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASEVDEGGGQLPEEI 312
                 :|||| ||:||  |  : ::  : |::| ||||:| :|  :     |  :
Db        237 -------INSSAKPIVYFIAGRDKSQRLWEPLRVVFQRALRDGAEPGDAASSTPNTV 286


RESULT 5
PCT-US93-08528-76
; Sequence 76, Application PC/TUS9308528
;  GENERAL INFORMATION:
;    APPLICANT:  New York University
;    TITLE OF INVENTION:  POLYPEPTIDES OF G-COUPLED PROTEIN
;    TITLE OF INVENTION:  RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
;    NUMBER OF SEQUENCES:  348
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/08528
;      FILING DATE:  09-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/943,236
;      FILING DATE:  10-SEP-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Townsend, Kevin G.
;      REGISTRATION NUMBER:  34,033
;      REFERENCE/DOCKET NUMBER:  MURPHY=2 PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  76:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  298 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
PCT-US93-08528-76

  Query Match             23.1%;  Score 384;  DB 5;  Length 298;
  Best Local Similarity   32.7%;  Pred. No. 8.4e-24;
  Matches   97;  Conservative   60;  Mismatches  102;  Indels   38;  Gaps   11;

Qy         36 IVSLVGLTGNAVVLWLLGCRMRRNAFSIYI--LNLAAADFLFLSGRLIYSLL---SFI-S 89
              :: | || || :||| |  ::|  ||||| |::|:|| ::|  : : :||   :|: |
Db          8 LLCLCGLVGNGLVLWFFGFSIKRTPFSIYIYFLHIASADGIYLFSKAVIALLNMGTFLGS 67

Qy         90 IPHTISKILYPVMMFSYFAGLSFLSAVSTERCLSVLWPIWYRCHRPTHLSAVVCVLLWAL 149
               |  : ::   :: ::|||:| | |:| |||:||::|:||   ||  ||| || ||| |
Db         68 FPDYVRRV-SRIVGLTFFAGVSLLPAISIERCVSVIFPMWYWRRRPKRLSAGVCALLWLL 126

Qy        150 SLLRSILEWMLCGFLFSGADSAWCQTSDFITVAWLIF-----LCVVLCGSSLVLLIRILC 204
              | | : :    |  |    |    |   :  |:|     : |: |    : |: : |
Db        127 SFLVTSIHNYFC-LLGHEASGTACLNMDISLLGILLFFLFCPIMVLPC----IALLHVEC 181

Qy        205 GSRKIPLT-RLYVTILLTVLVFLLCGLPFGIQFFLFLWI--------HVDREVLFCHVHL 255
               :|:   : :|   :|  | |||:  :  || :||| |:          |::  |
Db        182 RARRRQRSAKLNHVVLAIVSVFLVSSIYLGIDWFLF-WVFQIPAPFPEYVRDLCIC---- 236

Qy        256 VSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASEVDEGGGQLPEEI 312
                 :|||| ||:||  |  : ::  : |::| ||||:| :|  :     |  :
Db        237 -------INSSAKPIVYFIAGRDKSQRLWEPLRVVFQRALRDGAEPGDAASSTPNTV 286


RESULT 6
US-08-981-825-6
; Sequence 6, Application US/08981825
; Patent No. 6040426
;  GENERAL INFORMATION:
;    APPLICANT:  OGAWA, KAZUYAUKI
;    APPLICANT:  TANAKA, KAZUYA
;    APPLICANT:  NAGATA, KINYA
;    APPLICANT:  TAKANO, SYOICHI
;    TITLE OF INVENTION:  PROTEIN SPECIFIC TO HUMAN
;    TITLE OF INVENTION:  Th2, GENE (B19) ENCODING THE SAME, AND
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Knobbe, Martens, Olson & Bear
;      STREET:  620 Newport Center Drive  16th Floor
;      CITY:  Newport Beach
;      STATE:  CA
;      COUNTRY:  U.S.A.
;      ZIP:  92660
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
```

```
;      APPLICATION NUMBER:  US/08/981,825
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Altman, Daniel E
;      REGISTRATION NUMBER:  34,115
;      REFERENCE/DOCKET NUMBER:  MSHIM4.001APC
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  714-760-0404
;      TELEFAX:  714-760-9502
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  395 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
US-08-981-825-6

  Query Match             16.8%;  Score 279.5;  DB 3;  Length 395;
  Best Local Similarity   27.7%;  Pred. No. 2.9e-15;
  Matches   84;  Conservative   66;  Mismatches  114;  Indels   39;  Gaps   11;

Qy         36 IVSLVGLTGNAVVLWLLGCRMRRNAFSIYILNLAAADFLFLSGRLIYSLLSFISIPH--- 92
              : ||:||  | |:|:::|||||:   : ::|:|| :| |  :    ::   |::: |
Db         41 LASLLGLVENGVILFVVGCRMRQTVVTTWVLHLALSDLLASASLPFFTY--FLAVGHSWE 98

Qy         93 ---TISKILYPVMMFSYFAGLSFLSAVSTERCLSVLWPIWYRCHRPTHLSAVVCVLLWAL 149
                 |  |:   :   : ||    |||:| :||| |: |:| : ||   :  ||::||||
Db         99 LGTTFCKLHSSIFFLNMFASGFLLSAISLDRCLQVVRPVWAQNHRTVAAAHKVCLVLWAL 158

Qy        150 SLLRSILEW-----------MLCGF----LFSGAD-SAWCQT-------SDFITVAWLI 185
              ::| ::  :            ::| :    |  | |  | | :       | |: :|:|:
Db        159 AVLNTVPYFVFRDTISRLDGRIMCYYNVLLLNPGPDRDATCNSRQAALAVSKFL-LAFLV 217

Qy        186 FLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGIQFFLFLWIHVD 245
               | ::   |   : :|:    |: |  |     :   |  | ||  |: :  |    | :
Db        218 PLAII-ASSHAAVSLRLQHRGRRRP-GRFVRLVAAVVAAFALCWGPYHVFSLLEARAHAN 275

Qy        246 ---REVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASEVD 302
                 | :::  :  |:  |:  || |||::|       |: |::|: ||:  | | ||:
Db        276 PGLRPLVWRGLPFVT-SLAFFNSVANPVLYVLTCPDMLRKLRRSLRTVLESVLVDDSELG 334

Qy        303 EGG 305
               |
Db        335 GAG 337


RESULT 7
US-09-480-784-6
; Sequence 6, Application US/09480784
; Patent No. 6166186
;   GENERAL INFORMATION:
;        APPLICANT: OGAWA, KAZUYAUKI
;                   TANAKA, KAZUYA
;                   NAGATA, KINYA
;                   TAKANO, SYOICHI
;        TITLE OF INVENTION: PROTEIN SPECIFIC TO HUMAN
;                            Th2, GENE (B19) ENCODING THE SAME, AND
;        NUMBER OF SEQUENCES: 6
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Knobbe, Martens, Olson & Bear
;             STREET: 620 Newport Center Drive  16th Floor
;             CITY: Newport Beach
;             STATE: CA
;             COUNTRY: U.S.A.
;             ZIP: 92660
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: DOS
;             SOFTWARE: FastSEQ Version 1.5
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/480,784
;             FILING DATE: 10-Jan-2000
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/981,825
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Altman, Daniel E
;             REGISTRATION NUMBER: 34,115
;             REFERENCE/DOCKET NUMBER: MSHIM4.001APC
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 714-760-0404
;             TELEFAX: 714-760-9502
;             TELEX: <Unknown>
;   INFORMATION FOR SEQ ID NO: 6:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 395 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: unknown
;             TOPOLOGY: unknown
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-480-784-6

  Query Match             16.8%;  Score 279.5;  DB 4;  Length 395;
  Best Local Similarity   27.7%;  Pred. No. 2.9e-15;
  Matches   84;  Conservative   66;  Mismatches  114;  Indels   39;  Gaps   11;

Qy         36 IVSLVGLTGNAVVLWLLGCRMRRNAFSIYILNLAAADFLFLSGRLIYSLLSFISIPH--- 92
              : ||:||  | |:|:::|||||:   : ::|:|| :| |  :    ::   |::: |
Db         41 LASLLGLVENGVILFVVGCRMRQTVVTTWVLHLALSDLLASASLPFFTY--FLAVGHSWE 98

Qy         93 ---TISKILYPVMMFSYFAGLSFLSAVSTERCLSVLWPIWYRCHRPTHLSAVVCVLLWAL 149
                 |  |:   :   : ||    |||:| :||| |: |:| : ||   :  ||::||||
Db         99 LGTTFCKLHSSIFFLNMFASGFLLSAISLDRCLQVVRPVWAQNHRTVAAAHKVCLVLWAL 158

Qy        150 SLLRSILEW-----------MLCGF----LFSGAD-SAWCQT-------SDFITVAWLI 185
              ::| ::  :            ::| :    |  | |  | | :       | |: :|:|:
Db        159 AVLNTVPYFVFRDTISRLDGRIMCYYNVLLLNPGPDRDATCNSRQAALAVSKFL-LAFLV 217

Qy        186 FLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGIQFFLFLWIHVD 245
               | ::   |   : :|:    |: |  |     :   |  | ||  |: :  |    | :
Db        218 PLAII-ASSHAAVSLRLQHRGRRRP-GRFVRLVAAVVAAFALCWGPYHVFSLLEARAHAN 275

Qy        246 ---REVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASEVD 302
                 | :::  :  |:  |:  || |||::|       |: |::|: ||:  | | ||:
Db        276 PGLRPLVWRGLPFVT-SLAFFNSVANPVLYVLTCPDMLRKLRRSLRTVLESVLVDDSELG 334

Qy        303 EGG 305
               |
Db        335 GAG 337


RESULT 8
US-08-458-970A-9
; Sequence 9, Application US/08458970A
; Patent No. 5861272
; GENERAL INFORMATION:
;   APPLICANT:  LI, ET AL.
;   TITLE OF INVENTION:  C5a Receptor
;   NUMBER OF SEQUENCES:  11
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  CARELLA, BYRNE, BAIN, GILFILLAN,
;     ADDRESSEE:  CECCHI, STEWART & OLSTEIN
;     STREET:  6 BECKER FARM ROAD
;     CITY:  ROSELAND
;     STATE:  NEW JERSEY
;     COUNTRY:  USA
;     ZIP:  07068
```

```
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 INCH DISKETTE
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WORD PERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/458,970A
;      FILING DATE:  June 2, 1995
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/09234
;      FILING DATE:  16 AUG 1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MULLINS, J.G.
;      REGISTRATION NUMBER:  33,073
;      REFERENCE/DOCKET NUMBER:  325800-353
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-994-1700
;      TELEFAX:  201-994-1744
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  350 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-458-970A-9

  Query Match             15.1%;  Score 250;  DB 2;  Length 350;
  Best Local Similarity   26.6%;  Pred. No. 6.1e-13;
  Matches   89;  Conservative   56;  Mismatches  130;  Indels   60;  Gaps   10;

Qy          7 TLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNAFSIYIL 66
              |||   ||:: |  ||     |:| |:  :| |||: |||:|:|:     :|   :|: |
Db         19 TLDLN-TPVDKTSNTLRVPDILAL-VIFAVVFLVGVLGNALVVWVTAFEAKRTINAIWFL 76

Qy         67 NLAAADFL-FLSGRLIYSLLSFISIPH-----TISKILYPVMMFSYFAGLSFLSAVSTER 120
              ||| ||||  |:  ::::  | :   |          ||  ::: : :| :  |: :| :|
Db         77 NLAVADFLSCLALPILFT--SIVQHHHWPFGGAACSILPSLILLNMYASILLLATISADR 134

Qy        121 CLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQTSDFIT 180
              | |  ||| :  |   |: : | : | |:|| :|  ::              :
Db        135 FLLVFKPIWCQNFRGAGLAWIACAVAWGLALLLTIPSFL----------------YRV 176

Qy        181 VAWLIFLCVVLCG---------SSLVLLIRILCGS--RKIPLTRLYVTILL--------- 220
              |    |  ||||           | ::|:: |     : ||  |  |||
Db        177 VREEYFPPKVLCGVDYSHDKRRERAVAIVRLVLGFLWPLLTLTICYTFILLRTWSRRATR 236

Qy        221 ----------TVLVFLLCGLPFGIQFFLFLWIHVDREVLFCHVHLVSIFLS--ALNSSAN 268
                       |  | :  ||: :   :  ::             | |: :|   :|    |
Db        237 STKTLKVVVAVVASFFIFWLPYQVTGIMMSFLEPSSPTFLLLNKLDSLCVSFAYINCCIN 296

Qy        269 PIIYFFVGSFRQRQNRQNLKLVLQRALQDASEVDE 303
              ||||   |   | : |::|  :|:  | : | | |
Db        297 PIIYVVAGQGFQGRLRKSLPSLLRNVLTEESVVRE 331


RESULT 9
US-08-411-859-3
; Sequence 3, Application US/08411859
; Patent No. 5985600
;  GENERAL INFORMATION:
;    APPLICANT:  EVANS, CHRISTOPHER J.
;    APPLICANT:  KEITH JR., DUANE E.
;    APPLICANT:  EDWARDS, ROBERT H.
;    TITLE OF INVENTION:  EXPRESSION CLONING OF A DELTA OPIOID
;    TITLE OF INVENTION:  RECEPTOR, RELATED EXPRESSION SYSTEMS, AND RELATED
;    TITLE OF INVENTION:  PHARMACEUTICALS
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morrison & Foerster
;      STREET:  755 Page Mill Road
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94304-1018
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/411,859
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/929,200
;      FILING DATE:  13-AUG-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  LITHGOW, TIMOTHY J.
;      REGISTRATION NUMBER:  36,856
;      REFERENCE/DOCKET NUMBER:  22000-20526.00
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-813-5600
;      TELEFAX:  415-494-0792
;      TELEX:  706141
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  369 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-411-859-3

  Query Match             14.7%;  Score 243.5;  DB 2;  Length 369;
  Best Local Similarity   27.6%;  Pred. No. 2.2e-12;
  Matches   96;  Conservative   57;  Mismatches  126;  Indels   69;  Gaps   13;

Qy          3 PTISTLDTELTPINGTEETLCYKQTLSLTVLTCI---VSLVGLTGNAVVLW-LLGCRMRR 58
              |:   |:  | | ||: :|  |     |  ||| |   | :||| || :|:: :|     :
Db         16 PSPFDLNGSLGPSNGSNQTEPYYDMTSNAVLTFIYFVVCVVGLCGNTLVIYVILRYAKMK 75

Qy         59 NAFSIYILNLAAADFLFLSGRLIYSLLSFISIPH-----TISKILYPVMMFSYFAGLSFL 113
              :|||||| || ||: |  :  |   ::: |       | :::  |   : |  :  |
Db         76 TITNIYILNLAIADELFMLG--LPFLAMQVALVHWPFGKAICRVVMTVDGINQFTSIFCL 133

Qy        114 SAVSTERCLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWML------------ 160
              : :| :| |:|: ||      |   : :: | :| :||| ||  |:
Db        134 TVMSIDRYLAVVHPIKSAKWRRPRTAKMINVAVWGVSLL-VILPIMIYAGLRSNQWGRSS 192

Qy        161 CGFLFSGADSAWCQTSDFITVAWLIFLCV---VLCGSSLVLLIR-----ILCGSRKIPLT 212
              |   : |   ||   : ||  |:::   |   ::|   | ::|:     |  || |   :
Db        193 CTINWPGESGAW--YTGFIIYAFILGFLVPLTIICLCYLFIIIKVKSSGIRVGSSKRKKS 250

Qy        213 RLYVT--ILLTVLVFLLCGLPFGIQFFLFLWIHVDREVLFCHVHLVS------------ 257
                 ||  : : | ||: | ||     |::|             :|  ||
Db        251 EKKVTRMVSIVVAVFIFCWLP----FYIF-----------NVSSVSVAISPTPALKGMF 294

Qy        258 ---IFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASEVD 302
                : |:  || ||||:| |:    ::  :  | ||     :|    |
Db        295 DFVVILTYANSCANPILYAFLSDNFKKSFQNVLCLVKVSGAEDGERSD 342


RESULT 10
US-08-387-707-9
; Sequence 9, Application US/08387707
; Patent No. 6265563
;  GENERAL INFORMATION:
;    APPLICANT:  EVANS, CHRISTOPHER J.
;    APPLICANT:  KEITH, DUANE E.
;    TITLE OF INVENTION:  OPIOID RECEPTOR GENES
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  2000 Pennsylvania Avenue, N.W. Suite 5500
;      CITY:  Washington
```

;  STATE: DC
;  COUNTRY: USA
;  ZIP: 20006-1888
;  COMPUTER READABLE FORM:
;  MEDIUM TYPE: Floppy disk
;  COMPUTER: IBM PC compatible
;  OPERATING SYSTEM: PC-DOS/MS-DOS
;  SOFTWARE: PatentIn Release #1.0, Version #1.30
;  CURRENT APPLICATION DATA:
;  APPLICATION NUMBER: US/08/387,707
;  FILING DATE: 10-SEP-1995
;  CLASSIFICATION: 536
;  ATTORNEY/AGENT INFORMATION:
;  NAME: MURASHIGE, KATE H.
;  REGISTRATION NUMBER: 29,959
;  REFERENCE/DOCKET NUMBER: 22000-20526.20
;  TELECOMMUNICATION INFORMATION:
;  TELEPHONE: (202) 887-1500
;  TELEFAX: (202) 887-0763
;  INFORMATION FOR SEQ ID NO: 9:
;  SEQUENCE CHARACTERISTICS:
;  LENGTH: 369 amino acids
;  TYPE: amino acid
;  STRANDEDNESS: single
;  TOPOLOGY: linear

US-08-387-707-9

Query Match 14.7%; Score 243.5; DB 4; Length 369;
Best Local Similarity 27.6%; Pred. No. 2.2e-12;
Matches 96; Conservative 57; Mismatches 126; Indels 69; Gaps 13;

```
Qy          3 PTISTLDTELTPINGTEETLCYKQTLSLTVLTCI---VSLVGLTGNAVVLW-LLGCRMRR 58
              |:   |:  | | ||: :|  |    |  ||| |   | :||| || :|:: :|     :
Db         16 PSPFDLNGSLGPSNGSNQTEPYYDMTSNAVLTFIYFVVCVVGLCGNTLVIYVILRYAKMK 75

Qy         59 NAFSIYILNLAAADFLFLSGRLIYSLLSFISIPH-----TISKILYPVMMFSYFAGLSFL 113
                :|||||| || ||: |  :  |   ::: |      | :::  |   : |  :  |
Db         76 TITNIYILNLAIADELFMLG--LPFLAMQVALVHWPFGKAICRVVMTVDGINQFTSIFCL 133

Qy        114 SAVSTERCLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWML------------ 160
               : :| :|:|: ||      |    : :: | :| :|||  ||  |:
Db        134 TVMSIDRYLAVVHPIKSAKWRRPRTAKMINVAVWGVSLL-VILPIMIYAGLRSNQWGRSS 192

Qy        161 CGFLFSGADSAWCQTSDFITVAWLIFLCV---VLCGSSLVLLIR-----ILCGSRKIPLT 212
              |   : |   ||   : ||  |:::   |   ::|   | ::|:     |  || |   :
Db        193 CTINWPGESGAW--YTGFIIYAFILGFLVPLTIICLCYLFIIIKVKSSGIRVGSSKRKKS 250

Qy        213 RLYVT--ILLTVLVFLLCGLPFGIQFFLFLWIHVDREVLFCHVHLVS------------ 257
                ||  : : | ||: | ||    |::|            :|  ||
Db        251 EKKVTRMVSIVVAVFIFCWLP----FYIF-----------NVSSVSVAISPTPALKGMF 294

Qy        258 ---IFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASEVD 302
                : |:  || ||||:| |:    ::  :  | ||     :|    |
Db        295 DFVVILTYANSCANPILYAFLSDNFKKSFQNVLCLVKVSGAEDGERSD 342
```

RESULT 11
US-08-405-271A-9
; Sequence 9, Application US/08405271A
; Patent No. 6432652
; GENERAL INFORMATION:
; APPLICANT: EVANS, CHRISTOPHER J.
; APPLICANT: KEITH, DUANE E.
; TITLE OF INVENTION: OPIOID RECEPTOR GENES
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 PENNSYLVANIA AVENUE, NW, Suite 5500
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1888
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/405,271A
; FILING DATE: 14-MAR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 22000-20526.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030 MRSNFOERSWSH
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 369 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

US-08-405-271A-9

Query Match 14.7%; Score 243.5; DB 4; Length 369;
Best Local Similarity 27.6%; Pred. No. 2.2e-12;
Matches 96; Conservative 57; Mismatches 126; Indels 69; Gaps 13;

```
Qy          3 PTISTLDTELTPINGTEETLCYKQTLSLTVLTCI---VSLVGLTGNAVVLW-LLGCRMRR 58
              |:   |:  | | ||: :|  |    |  ||| |   | :||| || :|:: :|     :
Db         16 PSPFDLNGSLGPSNGSNQTEPYYDMTSNAVLTFIYFVVCVVGLCGNTLVIYVILRYAKMK 75

Qy         59 NAFSIYILNLAAADFLFLSGRLIYSLLSFISIPH-----TISKILYPVMMFSYFAGLSFL 113
                :|||||| || ||: |  :  |   ::: |      | :::  |   : |  :  |
Db         76 TITNIYILNLAIADELFMLG--LPFLAMQVALVHWPFGKAICRVVMTVDGINQFTSIFCL 133

Qy        114 SAVSTERCLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWML------------ 160
               : :| :|:|: ||      |    : :: | :| :|||  ||  |:
Db        134 TVMSIDRYLAVVHPIKSAKWRRPRTAKMINVAVWGVSLL-VILPIMIYAGLRSNQWGRSS 192

Qy        161 CGFLFSGADSAWCQTSDFITVAWLIFLCV---VLCGSSLVLLIR-----ILCGSRKIPLT 212
              |   : |   ||   : ||  |:::   |   ::|   | ::|:     |  || |   :
Db        193 CTINWPGESGAW--YTGFIIYAFILGFLVPLTIICLCYLFIIIKVKSSGIRVGSSKRKKS 250

Qy        213 RLYVT--ILLTVLVFLLCGLPFGIQFFLFLWIHVDREVLFCHVHLVS------------ 257
                ||  : : | ||: | ||    |::|            :|  ||
Db        251 EKKVTRMVSIVVAVFIFCWLP----FYIF-----------NVSSVSVAISPTPALKGMF 294

Qy        258 ---IFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASEVD 302
                : |:  || ||||:| |:    ::  :  | ||     :|    |
Db        295 DFVVILTYANSCANPILYAFLSDNFKKSFQNVLCLVKVSGAEDGERSD 342
```

RESULT 12
US-07-816-283-8
; Sequence 8, Application US/07816283
; Patent No. 5436155
; GENERAL INFORMATION:
; APPLICANT: Bell, Graeme I.
; APPLICANT: Yamada, Yuichiro
; APPLICANT: Seino, Susumu
; TITLE OF INVENTION: SOMATOSTATIN RECEPTORS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: PO Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/816,283
;      FILING DATE:  19911231
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McDaniel, C. Steven
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  713-787-1400
;      TELEFAX:  713-789-2679
;      TELEX:  79-0924
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  369 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-07-816-283-8

  Query Match             14.5%;  Score 241.5;  DB 1;  Length 369;
  Best Local Similarity   27.7%;  Pred. No. 3.1e-12;
  Matches   95;  Conservative   56;  Mismatches  123;  Indels   69;  Gaps   13;

Qy          8 LDTELTPINGTEETLCYKQTLSLTVLTCI---VSLVGLTGNAVVLW-LLGCRMRRNAFSI 63
              |:  | | ||:.:|  |     |  ||| |    | :||| || :|:: :|    :  :|
Db         21 LNGSLGPSNGSNQTEPYYDMTSNAVLTFIYFVVCVVGLCGNTLVIYVILRYAKMKTITNI 80

Qy         64 YILNLAAADFLFLSGRLIYSLLSFISIPH-----TISKILYPVMMFSYFAGLSFLSAVST 118
              |||||| || ||: |  :  |   ::: |      | :::  |   : |  :  |: :|
Db         81 YILNLAIADELFMLG--LPFLAMQVALVHWPFGKAICRVVMTVDGINQFTSIFCLTVMSI 138

Qy        119 ERCLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWML-------------CGFLF 165
              :| |:|: ||      |    : :: | :| :|||  ||  |:             |   :
Db        139 DRYLAVVHPIKSAKWRRPRTAKMINVAVWCVSLL-VILPIMIYAGLRSNQWGRSSCTINW 197

Qy        166 SGADSAWCQTSDFITVAWLIFLCV---VLCGSSLVLLIR-----ILCGSRKIPLTRLYVT 217
              |   ||   : ||  |:::   |   ::|   | ::|:     |  || |   :  ||
Db        198 PGESGAW--YTGFIIYAFILGFLVPLTIICLCYLFIIIKVKSSGIRVGSSKRKKSEKKVT 255

Qy        218 --ILLTVLVFLLCGLPFGIQFFLFLWIHVDREVLFCHVHLVS---------------IF 259
                : : | ||: | ||    |::|          :|  ||                :
Db        256 RMVSIVVAVFIFCWLP----FYIF-----------NVSSVSVAISPTPALKGMFDFVVI 299

Qy        260 LSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASEVD 302
              |:  || ||||:| |:    ::  :  | ||     :|    |
Db        300 LTYANSCANPILYAFLSDNFKKSFQNVLCLVKVSGTEDGERSD 342


RESULT 13
US-08-417-103-8
; Sequence 8, Application US/08417103
; Patent No. 5723299
;  GENERAL INFORMATION:
;    APPLICANT:  Bell, Graeme I.
;    APPLICANT:  Yamada, Yuichiro
;    APPLICANT:  Seino, Susumu
;    TITLE OF INVENTION:  SOMATOSTATIN RECEPTORS
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  United States of America
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/417,103
;      FILING DATE:  05-APR-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/816,283
;      FILING DATE:  01-DEC-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Wilson, Mark B.
;      REGISTRATION NUMBER:  37,259
;      REFERENCE/DOCKET NUMBER:  ARCD:144
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (512) 418-3000
;      TELEFAX:  (512) 474-7577
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  369 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-417-103-8

  Query Match             14.5%;  Score 241.5;  DB 1;  Length 369;
  Best Local Similarity   27.7%;  Pred. No. 3.1e-12;
  Matches   95;  Conservative   56;  Mismatches  123;  Indels   69;  Gaps   13;

Qy          8 LDTELTPINGTEETLCYKQTLSLTVLTCI---VSLVGLTGNAVVLW-LLGCRMRRNAFSI 63
              |:  | | ||: :|  |     |  ||| |    | :||| || :|:: :|    :  :|
Db         21 LNGSLGPSNGSNQTEPYYDMTSNAVLTFIYFVVCVVGLCGNTLVIYVILRYAKMKTITNI 80

Qy         64 YILNLAAADFLFLSGRLIYSLLSFISIPH-----TISKILYPVMMFSYFAGLSFLSAVST 118
              |||||| || ||: |  :  |   ::: |      | :::  |   : |  :  |: :|
Db         81 YILNLAIADELFMLG--LPFLAMQVALVHWPFGKAICRVVMTVDGINQFTSIFCLTVMSI 138

Qy        119 ERCLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWML-------------CGFLF 165
              :| |:|: ||      |    : :: | :| :|||  ||  |:             |   :
Db        139 DRYLAVVHPIKSAKWRRPRTAKMINVAVWCVSLL-VILPIMIYAGLRSNQWGRSSCTINW 197

Qy        166 SGADSAWCQTSDFITVAWLIFLCV---VLCGSSLVLLIR-----ILCGSRKIPLTRLYVT 217
              |   ||   : ||  |:::   |   ::|   | ::|:     |  || |   :  ||
Db        198 PGESGAW--YTGFIIYAFILGFLVPLTIICLCYLFIIIKVKSSGIRVGSSKRKKSEKKVT 255

Qy        218 --ILLTVLVFLLCGLPFGIQFFLFLWIHVDREVLFCHVHLVS---------------IF 259
                : : | ||: | ||    |::|          :|  ||                :
Db        256 RMVSIVVAVFIFCWLP----FYIF-----------NVSSVSVAISPTPALKGMFDFVVI 299

Qy        260 LSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASEVD 302
              |:  || ||||:| |:    ::  :  | ||     :|    |
Db        300 LTYANSCANPILYAFLSDNFKKSFQNVLCLVKVSGTEDGERSD 342


RESULT 14
US-08-120-601B-9
; Sequence 9, Application US/08120601B
; Patent No. 6235496
;  GENERAL INFORMATION:
;    APPLICANT:  Yu, Lei
;    TITLE OF INVENTION:  MU OPIOID RECEPTORS:  COMPOSOTIONS AND
;    TITLE OF INVENTION:  METHODS
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  USA
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/120,601B
;      FILING DATE:  13-SEP-1993
```

```
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Wilson, Mark B.
;      REGISTRATION NUMBER:  37,259
;      REFERENCE/DOCKET NUMBER:  INDA:002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  512/418-3000
;      TELEFAX:  512/474-7577
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  369 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
US-08-120-601B-9

  Query Match             14.5%;  Score 241.5;  DB 4;  Length 369;
  Best Local Similarity   27.6%;  Pred. No. 3.1e-12;
  Matches   96;  Conservative   58;  Mismatches  125;  Indels   69;  Gaps   14;

Qy          3 PTISTLDTELTPINGTEETLCYKQTLSLTVLTCI---VSLVGLTGNAVVLW-LLGCRMRR 58
              |:   |:  | | ||: :|  |    |  ||| |   | :||| || :|:: :|     :
Db         16 PSPFDLNGSLGPSNGSNQTEPYYDMTSNAVLTFIYFVVCVVGLCGNTLVIYVILRYAKMK 75

Qy         59 NAFSIYILNLAAADFLFLSGRLIYSLLSFISIPH-----TISKILYPVMMFSYFAGLSFL 113
                :|||||||| || ||: |  :  |   ::: |       | :::  |   :  |  :  |
Db         76 TITNIYILNLAIADELFMLG--LPFLAMQVALVHWPFGKAICRVVMTVDGINQFTSIFCL 133

Qy        114 SAVSTERCLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWML------------ 160
              : :| :| |:|: ||     |    : :: | :| :|||  ||  |:
Db        134 TVMSIDRYLAVVHPIKSAKWRRPRTAKMINVAVWGVSLL-VILPIMIYAGLRSNQWGRSS 192

Qy        161 CGFLFSGADSAWCQTSDFITVAWLI-FL----CVVLCGSSLVLLIR---ILCGSRKIPLT 212
              |  : |   ||   : ||  |::: ||     : ||   ::: ::   |  || |   :
Db        193 CTINWPGESGAW--YTGFIIYAFILGFLVPLTIICLCYLRIIIKVKSSGIRVGSSKRKKS 250

Qy        213 RLYVT--ILLTVLVFLLCGLPFGIQFFLFLWIHVDREVLFCHVHLVS------------ 257
                ||  : : | ||: | ||    |::|           :|  ||
Db        251 EKKVTRMVSIVVAVFIFCWLP----FYIF-----------NVSSVSVAISPTPALKGMF 294

Qy        258 ---IFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASEVD 302
                : |:  || ||||:| |:    :: :  | ||     :|    |
Db        295 DFVVILTYANSCANPILYAFLSDNFKKSFQNVLCLVKVSGAEDGERSD 342


RESULT 15
US-08-458-970A-10
; Sequence 10, Application US/08458970A
; Patent No. 5861272
;  GENERAL INFORMATION:
;    APPLICANT:  LI, ET AL.
;    TITLE OF INVENTION:  C5a Receptor
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CARELLA, BYRNE, BAIN, GILFILLAN,
;      ADDRESSEE:  CECCHI, STEWART & OLSTEIN
;      STREET:  6 BECKER FARM ROAD
;      CITY:  ROSELAND
;      STATE:  NEW JERSEY
;      COUNTRY:  USA
;      ZIP:  07068
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 INCH DISKETTE
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WORD PERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/458,970A
;      FILING DATE:  June 2, 1995
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/09234
;      FILING DATE:  16 AUG 1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MULLINS, J.G.
;      REGISTRATION NUMBER:  33,073
;      REFERENCE/DOCKET NUMBER:  325800-353
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-994-1700
;      TELEFAX:  201-994-1744
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  364 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-458-970A-10

  Query Match             14.2%;  Score 236;  DB 2;  Length 364;
  Best Local Similarity   24.8%;  Pred. No. 8.5e-12;
  Matches   84;  Conservative   57;  Mismatches  132;  Indels   66;  Gaps   13;

Qy         32 VLTCIVSLVGLTGNAVVLWLLGCRMRRNAFSIYILNLAAADFLFLSGRLIYSLLSFISIP 91
              ::  :  ::|: || :|:|: | ||:    :|  |||| ||| | |  | : : | :
Db         39 LIFAVTFVLGVLGNGLVIWVAGFRMKHTVTTISYLNLAIADFCFTS-TLPFYIASMVMGG 97

Qy         92 H-----TISKILYPVMMFSYFAGLSFLSAVSTERCLSVLWPIWYRCHRPTHLSAVVCVLL 146
              |       : | :| |:  : |  :  :: :: :||: || |:| : ||    |:  | ::
Db         98 HWPFGWFMCKFIYTVIDINLFGSVFLIALIALDRCICVLHPVWAQNHRTVSLAKKVIIVP 157

Qy        147 WALSLLRSILEWMLCGFLFS--------------GADSAWCQTSDF------------I 179
              |              :| || :                |     | | ||          :
Db        158 W------------ICAFLLTLPVIIRLTTVPNSRLGPGKTAC-TFDFSPWTKDPVEKRKV 204

Qy        180 TVAWLIFLCVV--LCGSSLVLLIRILC---------GSRKIPLTRLYVTILLTVLVFLLC 228
              | |   ::  : | |  : |               |  :|    :  | | ||
Db        205 AVTMLTVRGIIRFIIGFSTPMSIVAICYGLITTKIHRQGLIKSSRPLRVLSFVVAAFFLC 264

Qy        229 GLPFGIQFFLFLWIHVDREVLFCHVHLVSIF-----LSALNSSANPIIYFFVG-SFRQRQ 282
              || :   |   | |   :      :|:       |:  ||  ||::| |:|  ||:|
Db        265 WCPFQV-VALISTIQVRERLKNMTPGIVTALKITSPLAFFNSCLNPMLYVFMGQDFRERL 323

Qy        283 NRQNLKLVLQRAL-QDASEVDEGGGQLPEEILELSGSRL 320
              :|   |:|| :|:::  : |  |      || : |
Db        324 -IHSLPASLERALTEDSAQTSDTGTNLGTNSTSLSENTL 361


Search completed:  July  2, 2003, 19:23:15
Job time :  15 secs
```

GenCore version 5.1.6


OM protein - protein search, using sw model

Run on: July 2, 2003, 19:16:35 ; Search time 80 Seconds
(without alignments)
829.340 Million cell updates/sec

Title: US-09-849-869A-16
Perfect score: 1661
Sequence: 1 MDPTISTLDTELTPINGTEE.....EGGGQLPEEILELSGSRLEQ 322

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 1661 | 100.0 | 322 | 4 | Q96LB2 | Q961b2 homo sapien |
| 2 | 1660 | 99.9 | 322 | 4 | Q8TDD8 | Q8tdd8 homo sapien |
| 3 | 1624 | 97.8 | 322 | 4 | Q8TDD9 | Q8tdd9 homo sapien |
| 4 | 1383 | 83.3 | 322 | 4 | Q96LB0 | Q961b0 homo sapien |
| 5 | 1373 | 82.7 | 322 | 4 | Q8TDE1 | Q8tde1 homo sapien |
| 6 | 1313 | 79.0 | 322 | 4 | Q8TDE0 | Q8tde0 homo sapien |
| 7 | 1309 | 78.8 | 322 | 4 | Q8TDD7 | Q8tdd7 homo sapien |
| 8 | 1294 | 77.9 | 322 | 4 | Q96LA9 | Q961a9 homo sapien |
| 9 | 1287 | 77.5 | 322 | 4 | Q8TDD6 | Q8tdd6 homo sapien |
| 10 | 1038.5 | 62.5 | 330 | 4 | Q96LB1 | Q961b1 homo sapien |
| 11 | 847 | 51.0 | 337 | 11 | Q8R4G1 | Q8r4g1 rattus norv |
| 12 | 821 | 49.4 | 338 | 11 | Q91ZC2 | Q91zc2 mus musculu |
| 13 | 769.5 | 46.3 | 338 | 11 | Q91ZC3 | Q91zc3 mus musculu |
| 14 | 761.5 | 45.8 | 331 | 11 | Q91YB7 | Q91yb7 rattus norv |
| 15 | 756.5 | 45.5 | 304 | 11 | Q91WW5 | Q91ww5 mus musculu |
| 16 | 735 | 44.3 | 302 | 11 | Q91WW3 | Q91ww3 mus musculu |
| 17 | 731.5 | 44.0 | 301 | 11 | Q91ZC6 | Q91zc6 mus musculu |
| 18 | 731.5 | 44.0 | 321 | 11 | Q91ZC0 | Q91zc0 mus musculu |
| 19 | 716.5 | 43.1 | 305 | 11 | Q91ZC5 | Q91zc5 mus musculu |
| 20 | 704.5 | 42.4 | 305 | 11 | Q91ZC4 | Q91zc4 mus musculu |
| 21 | 692.5 | 41.7 | 322 | 11 | Q91ZB9 | Q91zb9 mus musculu |
| 22 | 690.5 | 41.6 | 305 | 11 | Q91WW4 | Q91ww4 mus musculu |
| 23 | 687.5 | 41.4 | 313 | 11 | Q91WW2 | Q91ww2 mus musculu |
| 24 | 675 | 40.6 | 287 | 11 | Q91ZC7 | Q91zc7 mus musculu |
| 25 | 634 | 38.2 | 312 | 11 | Q91ZC1 | Q91zc1 mus musculu |
| 26 | 538.5 | 32.4 | 321 | 4 | Q8TDS7 | Q8tds7 homo sapien |
| 27 | 535.5 | 32.2 | 321 | 11 | Q91ZB8 | Q91zb8 mus musculu |
| 28 | 484 | 29.1 | 310 | 11 | Q91ZB7 | Q91zb7 mus musculu |
| 29 | 480.5 | 28.9 | 324 | 11 | Q9D5Q2 | Q9d5q2 mus musculu |
| 30 | 451 | 27.2 | 343 | 4 | Q96AM1 | Q96am1 homo sapien |
| 31 | 442.5 | 26.6 | 321 | 11 | Q99MT8 | Q99mt8 mus musculu |
| 32 | 416 | 25.0 | 319 | 11 | Q91ZB6 | Q91zb6 mus musculu |
| 33 | 416 | 25.0 | 343 | 11 | Q8VCJ6 | Q8vcj6 mus musculu |
| 34 | 349 | 21.0 | 289 | 11 | Q91ZB5 | Q91zb5 mus musculu |
| 35 | 253 | 15.2 | 390 | 13 | Q8QGQ4 | Q8qgq4 carassius a |
| 36 | 250 | 15.1 | 351 | 11 | O88536 | O88536 mus musculu |
| 37 | 248.5 | 15.0 | 343 | 11 | O88537 | O88537 mus musculu |
| 38 | 247.5 | 14.9 | 355 | 6 | Q95LH1 | Q95lh1 macaca fasc |
| 39 | 241.5 | 14.5 | 475 | 11 | Q9QWG9 | Q9qwg9 cavia porce |
| 40 | 238 | 14.3 | 370 | 13 | Q8UWL5 | Q8uwl5 fugu rubrip |
| 41 | 237.5 | 14.3 | 346 | 11 | Q91Y73 | Q91y73 mus musculu |
| 42 | 234.5 | 14.1 | 323 | 11 | O88538 | O88538 mus musculu |
| 43 | 233.5 | 14.1 | 380 | 13 | Q9DGQ6 | Q9dgq6 carassius a |
| 44 | 229.5 | 13.8 | 470 | 4 | Q8TDT2 | Q8tdt2 homo sapien |
| 45 | 224 | 13.5 | 385 | 11 | Q9JK40 | Q9jk40 mus musculu |

ALIGNMENTS

RESULT 1
Q96LB2
ID Q96LB2 PRELIMINARY; PRT; 322 AA.
AC Q96LB2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE G protein-coupled receptor (Putative G-protein coupled receptor).
GN MRGX1 OR GPCR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21435808; PubMed=11551509;
RA Dong X., Han S.-K., Zylka M.J., Simon M.I., Anderson D.J.;
RT "A Diverse Family of GPCRs Expressed in Specific Subsets of
RT Nociceptive Somatosensory Neurons.";
RL Cell 106:619-632(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Takeda S., Kadowaki S., Haga T., Takaesu H., Mitaku S.;
RT "Identification of G protein-coupled receptor genes from the human
RT genome sequence.";
RL Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY042213; AAK91804.1; -.
DR EMBL; AB083628; BAB89341.1; -.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 322 AA; 36250 MW; C7F3A9F4418E8AD1 CRC64;

Query Match 100.0%; Score 1661; DB 4; Length 322;
Best Local Similarity 100.0%; Pred. No. 6e-143;
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MDPTISTLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDPTISTLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNA 60

Qy         61 FSIYILNLAAADFLFLSGRLIYSLLSFISIPHTISKILYPVMMFSYFAGLSFLSAVSTER 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 FSIYILNLAAADFLFLSGRLIYSLLSFISIPHTISKILYPVMMFSYFAGLSFLSAVSTER 120

Qy        121 CLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQTSDFIT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQTSDFIT 180

Qy        181 VAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGIQFFLFL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGIQFFLFL 240

Qy        241 WIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 WIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASE 300

Qy        301 VDEGGGQLPEEILELSGSRLEQ 322
              ||||||||||||||||||||||
Db        301 VDEGGGQLPEEILELSGSRLEQ 322
```

RESULT 2
Q8TDD8
ID Q8TDD8 PRELIMINARY; PRT; 322 AA.
AC Q8TDD8;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE G protein-coupled receptor SNSR4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21853733; PubMed=11850634;
RA Lembo P.M., Grazzini E., Groblewski T., O'Donnell D., Roy M.-O.,
RA Zhang J., Hoffert C., Cao J., Schmidt R., Pelletier M., Labarre M.,
RA Gosselin M., Fortin Y., Banville D., Shen S., Strom P., Payza K.,
RA Dray A., Walker P., Ahmad S.;
RT "Proenkephalin A gene products activate a new family of sensory
RT neuron-specific GPCRs.";
RL Nat. Neurosci. 5:201-209(2002).
DR EMBL; AF474990; AAL86881.1; -.
KW Receptor.
SQ SEQUENCE 322 AA; 36236 MW; C301BC174BB01D72 CRC64;

Query Match 99.9%; Score 1660; DB 4; Length 322;
Best Local Similarity 99.7%; Pred. No. 7.3e-143;
Matches 321; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MDPTISTLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNA 60
              ||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDPTVSTLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNA 60

Qy         61 FSIYILNLAAADFLFLSGRLIYSLLSFISIPHTISKILYPVMMFSYFAGLSFLSAVSTER 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 FSIYILNLAAADFLFLSGRLIYSLLSFISIPHTISKILYPVMMFSYFAGLSFLSAVSTER 120

Qy        121 CLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQTSDFIT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQTSDFIT 180

Qy        181 VAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGIQFFLFL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGIQFFLFL 240

Qy        241 WIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 WIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASE 300

Qy        301 VDEGGGQLPEEILELSGSRLEQ 322
              ||||||||||||||||||||||
Db        301 VDEGGGQLPEEILELSGSRLEQ 322
```

RESULT 3
Q8TDD9
ID Q8TDD9 PRELIMINARY; PRT; 322 AA.
AC Q8TDD9;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE G protein-coupled receptor SNSR3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21853733; PubMed=11850634;
RA Lembo P.M., Grazzini E., Groblewski T., O'Donnell D., Roy M.-O.,
RA Zhang J., Hoffert C., Cao J., Schmidt R., Pelletier M., Labarre M.,
RA Gosselin M., Fortin Y., Banville D., Shen S., Strom P., Payza K.,
RA Dray A., Walker P., Ahmad S.;
RT "Proenkephalin A gene products activate a new family of sensory
RT neuron-specific GPCRs.";
RL Nat. Neurosci. 5:201-209(2002).
DR EMBL; AF474989; AAL86880.1; -.
KW Receptor.
SQ SEQUENCE 322 AA; 36300 MW; 6A1BB33D6EC7077B CRC64;

Query Match 97.8%; Score 1624; DB 4; Length 322;
Best Local Similarity 97.2%; Pred. No. 1.4e-139;
Matches 313; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

```
Qy          1 MDPTISTLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNA 60
              ||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDPTVSTLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNA 60

Qy         61 FSIYILNLAAADFLFLSGRLIYSLLSFISIPHTISKILYPVMMFSYFAGLSFLSAVSTER 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||:|
Db         61 FSIYILNLAAADFLFLSGRLIYSLLSFISIPHTISKILYPVMMFSYFAGLNFLSAVSTDR 120

Qy        121 CLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQTSDFIT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQTSDFIT 180

Qy        181 VAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGIQFFLFL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGIQFFLFL 240

Qy        241 WIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASE 300
              |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||  |
Db        241 WIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSLRQRQNRQNLKLVLQRALQDTPE 300

Qy        301 VDEGGGQLPEEILELSGSRLEQ 322
              |||||| ||:| ||||||||||
Db        301 VDEGGGWLPQETLELSGSRLEQ 322
```

RESULT 4
Q96LB0
ID Q96LB0 PRELIMINARY; PRT; 322 AA.
AC Q96LB0;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE G protein-coupled receptor.
GN MRGX3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21435808; PubMed=11551509;
RA   Dong X., Han S.-K., Zylka M.J., Simon M.I., Anderson D.J.;
RT   "A Diverse Family of GPCRs Expressed in Specific Subsets of
RT   Nociceptive Somatosensory Neurons.";
RL   Cell 106:619-632(2001).
DR   EMBL; AY042215; AAK91806.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   322 AA;  36484 MW;  253B1BEF0CB4EB74 CRC64;

  Query Match             83.3%;  Score 1383;  DB 4;  Length 322;
  Best Local Similarity   83.5%;  Pred. No. 1.1e-117;
  Matches  269;  Conservative   16;  Mismatches   37;  Indels    0;  Gaps    0;

Qy          1 MDPTISTLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNA 60
              || ||  | |||||||| ||| ||||||| | |||||||| ||||||||||||||||||
Db          1 MDSTIPVLGTELTPINGREETPCYKQTLSFTGLTCIVSLVALTGNAVVLWLLGCRMRRNA 60

Qy         61 FSIYILNLAAADFLFLSGRLIYSLLSFISIPHTISKILYPVMMFSYFAGLSFLSAVSTER 120
              ||||||| |||||||| :| | |  |:| | ||||| ||| | || ||| |||:||||
Db         61 VSIYILNLVAADFLFLSGHIICSPLRLINIRHPISKILSPVMTFPYFIGLSMLSAISTER 120

Qy        121 CLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQTSDFIT 180
              |||:|||||| | || :||:|:|||||||||||||||| | ||||||| ||:||||||
Db        121 CLSILWPIWYHCRRPRYLSSVMCVLLWALSLLRSILEWMFCDFLFSGADSVWCETSDFIT 180

Qy        181 VAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGIQFFLFL 240
              :|||:||||||||||||:|||||||:|||||||||||||||||||||||||||||: ||
Db        181 IAWLVFLCVVLCGSSLVLLVRILCGSRKMPLTRLYVTILLTVLVFLLCGLPFGIQWALFS 240

Qy        241 WIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASE 300
              ||:| :|||||||||||||||||||||||||||||||||||||||||||||||||||  |
Db        241 RIHLDWKVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDTPE 300

Qy        301 VDEGGGQLPEEILELSGSRLEQ 322
              |||||| ||:| ||||||||||
Db        301 VDEGGGWLPQETLELSGSRLEQ 322


RESULT 5
Q8TDE1
ID   Q8TDE1      PRELIMINARY;      PRT;   322 AA.
AC   Q8TDE1;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   G protein-coupled receptor SNSR1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21853733; PubMed=11850634;
RA   Lembo P.M., Grazzini E., Groblewski T., O'Donnell D., Roy M.-O.,
RA   Zhang J., Hoffert C., Cao J., Schmidt R., Pelletier M., Labarre M.,
RA   Gosselin M., Fortin Y., Banville D., Shen S., Strom P., Payza K.,
RA   Dray A., Walker P., Ahmad S.;
RT   "Proenkephalin A gene products activate a new family of sensory
RT   neuron-specific GPCRs.";
RL   Nat. Neurosci. 5:201-209(2002).
DR   EMBL; AF474987; AAL86878.1; -.
KW   Receptor.
SQ   SEQUENCE   322 AA;  36448 MW;  30A7F306F67C64D7 CRC64;

  Query Match             82.7%;  Score 1373;  DB 4;  Length 322;
  Best Local Similarity   82.0%;  Pred. No. 8.6e-117;
  Matches  264;  Conservative   22;  Mismatches   36;  Indels    0;  Gaps    0;

Qy          1 MDPTISTLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNA 60
              |||||  | |:||||||| ||| || |||| | ||||:||| ||||||||||||||||||
Db          1 MDPTIPVLGTKLTPINGREETPCYNQTLSFTGLTCIISLVALTGNAVVLWLLGCRMRRNA 60

Qy         61 FSIYILNLAAADFLFLSGRLIYSLLSFISIPHTISKILYPVMMFSYFAGLSFLSAVSTER 120
              ||||||| ||:|||||| :|:| |  |:| | ||||| ||| | || ||| |||:||||
Db         61 VSIYILNLVAANFLFLSGHIIFSPLPLINIRHPISKILSPVMTFPYFIGLSMLSAISTER 120

Qy        121 CLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQTSDFIT 180
              |||:|||||| | || :||:|:|||||||||||||||| | ||||||:| ||:||||||
Db        121 CLSILWPIWYHCRRPRYLSSVMCVLLWALSLLRSILEWMFCDFLFSGANSVWCETSDFIT 180

Qy        181 VAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGIQFFLFL 240
              :|||:|||||||||||||:|||||||:|||||||||||||||||||||||||||||: ||
Db        181 IAWLVFLCVVLCGSSLVLLVRILCGSRKMPLTRLYVTILLTVLVFLLCGLPFGIQWALFS 240

Qy        241 WIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASE 300
              ||:| :|||||||||||||||||||||||||||||||||||||||||||||||||||  |
Db        241 RIHLDWKVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDTPE 300

Qy        301 VDEGGGQLPEEILELSGSRLEQ 322
              |||||| ||:| |||||||:|||
Db        301 VDEGGGWLPQETLELSGSKLEQ 322


RESULT 6
Q8TDE0
ID   Q8TDE0      PRELIMINARY;      PRT;   322 AA.
AC   Q8TDE0;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   G protein-coupled receptor SNSR2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21853733; PubMed=11850634;
RA   Lembo P.M., Grazzini E., Groblewski T., O'Donnell D., Roy M.-O.,
RA   Zhang J., Hoffert C., Cao J., Schmidt R., Pelletier M., Labarre M.,
RA   Gosselin M., Fortin Y., Banville D., Shen S., Strom P., Payza K.,
RA   Dray A., Walker P., Ahmad S.;
RT   "Proenkephalin A gene products activate a new family of sensory
RT   neuron-specific GPCRs.";
RL   Nat. Neurosci. 5:201-209(2002).
DR   EMBL; AF474988; AAL86879.1; -.
KW   Receptor.
SQ   SEQUENCE   322 AA;  36551 MW;  CE42431FD3B4611B CRC64;

  Query Match             79.0%;  Score 1313;  DB 4;  Length 322;
  Best Local Similarity   80.1%;  Pred. No. 2.4e-111;
  Matches  257;  Conservative   21;  Mismatches   43;  Indels    0;  Gaps    0;

Qy          1 MDPTISTLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNA 60
              ||||:  | |||||||| ||| ||||||| | |||||||| ||||||||||||||||||
Db          1 MDPTVPVLGTELTPINGREETPCYKQTLSFTGLTCIVSLVALTGNAVVLWLLGCRMRRNA 60

Qy         61 FSIYILNLAAADFLFLSGRLIYSLLSFISIPHTISKILYPVMMFSYFAGLSFLSAVSTER 120
              ||||||| ||||||||| :| | |  |:| | ||||| ||| | || ||| |:|:||||
Db         61 VSIYILNLVAADFLFLSGHIICSPLRLINISHPISKILSPVMTFPYFIGLSMLNAISTER 120

Qy        121 CLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQTSDFIT 180
              |||:|||||| | || :||:|:|||||| |||||||||| | ||||||||  |:||||||
Db        121 CLSILWPIWYHCRRPRYLSSVMCVLLWAPSLLRSILEWMFCDFLFSGADSVRCETSDFIT 180

Qy        181 VAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGIQFFLFL 240
              :|||:|| |||||||||||:||||||||:|||||||||||||||||||||||||||||: ||
Db        181 IAWLVFLRVVLCGSSLVLLVRILCGSRKMPLTRLYVTILLTVLVFLLCGLPFGIQWALFS 240
```

```
Qy        241 WIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASE 300
              ||:| :||||||||||||||||||||||||||:||||| |||: ||||||| |||  |
Db        241 RIHLDWKVLFCHVHLVSIFLSALNSSANPIIYFFMGSFRQLQNRKTLKLVLQRDLQDTPE 300

Qy        301 VDEGGGQLPEEILELSGSRLE 321
              |||||  ||:| ||||||:||
Db        301 VDEGGWWLPQETLELSGSKLE 321


RESULT 7
Q8TDD7
ID   Q8TDD7      PRELIMINARY;      PRT;   322 AA.
AC   Q8TDD7;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   G protein-coupled receptor SNSR5.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21853733; PubMed=11850634;
RA   Lembo P.M., Grazzini E., Groblewski T., O'Donnell D., Roy M.-O.,
RA   Zhang J., Hoffert C., Cao J., Schmidt R., Pelletier M., Labarre M.,
RA   Gosselin M., Fortin Y., Banville D., Shen S., Strom P., Payza K.,
RA   Dray A., Walker P., Ahmad S.;
RT   "Proenkephalin A gene products activate a new family of sensory
RT   neuron-specific GPCRs.";
RL   Nat. Neurosci. 5:201-209(2002).
DR   EMBL; AF474991; AAL86882.1; -.
KW   Receptor.
SQ   SEQUENCE   322 AA;  36423 MW;  3D6FFB4B5DDFDD90 CRC64;

  Query Match             78.8%;  Score 1309;  DB 4;  Length 322;
  Best Local Similarity   80.3%;  Pred. No. 5.6e-111;
  Matches  257;  Conservative   20;  Mismatches   43;  Indels    0;  Gaps    0;

Qy          1 MDPTISTLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNA 60
              ||||:  | |:|||||| ||| ||||||| ||||||:|||||||||||||||||||||||
Db          1 MDPTVPVLGTKLTPINGREETPCYKQTLSFTVLTCIISLVGLTGNAVVLWLLGCRMRRNA 60

Qy         61 FSIYILNLAAADFLFLSGRLIYSLLSFISIPHTISKILYPVMMFSYFAGLSFLSAVSTER 120
              |||||||||||||||| ::| | |  |:| | | |||  || | || ||| |||:||||
Db         61 VSIYILNLAAADFLFLSFQIIRSPLRLINISHLIRKILVSVMTFPYFTGLSMLSAISTER 120

Qy        121 CLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQTSDFIT 180
              |||||||||||| |||||||||||||| |||| |:|||  | ||||||||:||:|||||
Db        121 CLSVLWPIWYRCRRPTHLSAVVCVLLWGLSLLFSMLEWRFCDFLFSGADSSWCETSDFIP 180

Qy        181 VAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGIQFFLFL 240
              |||||||||||| ||||||:||||||||:||||||||||||||||||||||||    |
Db        181 VAWLIFLCVVLCVSSLVLLVRILCGSRKMPLTRLYVTILLTVLVFLLCGLPFGILGALIY 240

Qy        241 WIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASE 300
              :|:: |||:|||:|| : ||:||||||||||||||||||||||||||||||||||||  |
Db        241 RMHLNLEVLYCHVYLVCMSLSSLNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDKPE 300

Qy        301 VDEGGGQLPEEILELSGSRL 320
              ||:| |||||| |||||| ||
Db        301 VDKGEGQLPEESLELSGRRL 320


RESULT 8
Q96LA9
ID   Q96LA9      PRELIMINARY;      PRT;   322 AA.
AC   Q96LA9;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   G protein-coupled receptor.
GN   MRGX4.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21435808; PubMed=11551509;
RA   Dong X., Han S.-K., Zylka M.J., Simon M.I., Anderson D.J.;
RT   "A Diverse Family of GPCRs Expressed in Specific Subsets of
RT   Nociceptive Somatosensory Neurons.";
RL   Cell 106:619-632(2001).
DR   EMBL; AY042216; AAK91807.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   322 AA;  36434 MW;  7CA676F8BD390A31 CRC64;

  Query Match             77.9%;  Score 1294;  DB 4;  Length 322;
  Best Local Similarity   79.7%;  Pred. No. 1.3e-109;
  Matches  255;  Conservative   20;  Mismatches   45;  Indels    0;  Gaps    0;

Qy          1 MDPTISTLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNA 60
              ||||:    |:|||||| ||| || |||| ||||||:|||||||||||||||| ||||||
Db          1 MDPTVPVFGTKLTPINGREETPCYNQTLSFTVLTCIISLVGLTGNAVVLWLLGYRMRRNA 60

Qy         61 FSIYILNLAAADFLFLSGRLIYSLLSFISIPHTISKILYPVMMFSYFAGLSFLSAVSTER 120
              |||||||||||||||| ::| | |  |:| | | |||  || | || ||| |||:||||
Db         61 VSIYILNLAAADFLFLSFQIIRSPLRLINISHLIRKILVSVMTFPYFTGLSMLSAISTER 120

Qy        121 CLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQTSDFIT 180
              |||||||||||| |||||||||||||| |||| |:|||  | ||||||||:||:|||||
Db        121 CLSVLWPIWYRCRRPTHLSAVVCVLLWGLSLLFSMLEWRFCDFLFSGADSSWCETSDFIP 180

Qy        181 VAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGILGALIY 240
              |||||||||||| ||||||:||||||||:||||||||||||||||||||||||    |
Db        181 VAWLIFLCVVLCVSSLVLLVRILCGSRKMPLTRLYVTILLTVLVFLLCGLPFGILGALIY 240

Qy        241 WIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDKPE 300
              :|:: |||:|||:|| : ||:||||||||||||||||||||||||||||||||||||  |
Db        241 RMHLNLEVLYCHVYLVCMSLSSLNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDKPE 300

Qy        301 VDKGEGQLPEESLELSGSRL 320
              ||:| |||||| |||||| ||
Db        301 VDKGEGQLPEESLELSGSRL 320


RESULT 9
Q8TDD6
ID   Q8TDD6      PRELIMINARY;      PRT;   322 AA.
AC   Q8TDD6;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   G protein-coupled receptor SNSR6.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21853733; PubMed=11850634;
RA   Lembo P.M., Grazzini E., Groblewski T., O'Donnell D., Roy M.-O.,
RA   Zhang J., Hoffert C., Cao J., Schmidt R., Pelletier M., Labarre M.,
RA   Gosselin M., Fortin Y., Banville D., Shen S., Strom P., Payza K.,
RA   Dray A., Walker P., Ahmad S.;
RT   "Proenkephalin A gene products activate a new family of sensory
RT   neuron-specific GPCRs.";
RL   Nat. Neurosci. 5:201-209(2002).
DR   EMBL; AF474992; AAL86883.1; -.
KW   Receptor.
```

```
SQ   SEQUENCE   322 AA;  36434 MW;  7CA66E58D70548BA CRC64;

  Query Match             77.5%;  Score 1287;  DB 4;  Length 322;
  Best Local Similarity   79.1%;  Pred. No. 5.5e-109;
  Matches  253;  Conservative   21;  Mismatches   46;  Indels    0;  Gaps    0;

Qy          1 MDPTISTLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNA 60
              ||||:    |:|||||| ||| || |||| ||||||:|||||||||||||||| ||||||
Db          1 MDPTVPVFGTKLTPINGREETPCYNQTLSFTVLTCIISLVGLTGNAVVLWLLGYRMRRNA 60

Qy         61 FSIYILNLAAADFLFLSGRLIYSLLSFISIPHTISKILYPVMMFSYFAGLSFLSAVSTER 120
              |||||||||||||||| ::| | |  |:| | | |||  || | || ||| |||:|||| 
Db         61 VSIYILNLAAADFLFLSFQIIRSPLRLINISHLIRKILVSVMTFPYFTGLSMLSAISTER 120

Qy        121 CLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQTSDFIT 180
              |||||||||||||| |||||||||||||| |||| |:|||  | ||||||||:||:|||||
Db        121 CLSVLWPIWYRCRRPTHLSAVVCVLLWGLSLLFSMLEWRFCDFLFSGADSSWCETSDFIP 180

Qy        181 VAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGIQFFLFL 240
              | |||||||||| |||||||:||||||||:|||||||||||||||||||||||| | |
Db        181 VVWLIFLCVVLCVSSLVLLVRILCGSRKMPLTRLYVTILLTVLVFLLCGLPFGILGALIY 240

Qy        241 WIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASE 300
              :|:: |||:|||:|| : ||:||||||||||||||||||||||||||||||||||  |
Db        241 RMHLNLEVLYCHVYLVCMSLSSLNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDKPE 300

Qy        301 VDEGGGQLPEEILELSGSRL 320
              ||:| |||||| ||||||:|
Db        301 VDKGEGQLPEESLELSGSKL 320


RESULT 10
Q96LB1
ID   Q96LB1      PRELIMINARY;      PRT;   330 AA.
AC   Q96LB1;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   G protein-coupled receptor (Putative G-protein coupled receptor).
GN   MRGX2 OR GPCR.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21435808; PubMed=11551509;
RA   Dong X., Han S.-K., Zylka M.J., Simon M.I., Anderson D.J.;
RT   "A Diverse Family of GPCRs Expressed in Specific Subsets of
RT   Nociceptive Somatosensory Neurons.";
RL   Cell 106:619-632(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Takeda S., Kadowaki S., Haga T., Takaesu H., Mitaku S.;
RT   "Identification of G protein-coupled receptor genes from the human
RT   genome sequence.";
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY042214; AAK91805.1; -.
DR   EMBL; AB083626; BAB89339.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   330 AA;  37099 MW;  0B328FD78B1DF6BE CRC64;

  Query Match             62.5%;  Score 1038.5;  DB 4;  Length 330;
  Best Local Similarity   64.7%;  Pred. No. 2.1e-86;
  Matches  213;  Conservative   27;  Mismatches   80;  Indels    9;  Gaps    3;

Qy          1 MDPTISTLDTELTPINGTEET---LCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMR 57
              ||||     || | :|| ::    || |:||     |   ::||||| ||  |||||| |||
Db          1 MDPTTPAWGTESTTVNGNDQALLLLCGKETLIPVFLILFIALVGLVGNGFVLWLLGFRMR 60

Qy         58 RNAFSIYILNLAAADFLFLSGRLIYSLL----SFISIPHTISKILYPVMMFSYFAGLSFL 113
              |||||:|:|:|| |||||  ::|  |:      | ||           ||  :| |||| |
Db         61 RNAFSVYVLSLAGADFLFLCFQIINCLVYLSNFFCSISINFPSFFTTVMTCAYLAGLSML 120

Qy        114 SAVSTERCLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWC 173
              | |||||||||||||||||| || ||||||||||||||||| ||||   ||||||  || ||
Db        121 STVSTERCLSVLWPIWYRCRRPRHLSAVVCVLLWALSLLLSILEGKFCGFLFSDGDSGWC 180

Qy        174 QTSDFITVAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFG 233
              || |||| |||||| :||||||| ||:||||||| :||||||:|||||||||||||||||
Db        181 QTFDFITAAWLIFLFMVLCGSSLALLVRILCGSRGLPLTRLYLTILLTVLVFLLCGLPFG 240

Qy        234 IQFFLFLWIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQN--LKLVL 291
              ||:|| |||  | :|||||:| ||: ||:|||||||||||||||::   |   ||| |
Db        241 IQWFLILWIWKDSDVLFCHIHPVSVVLSSLNSSANPIIYFFVGSFRKQWRLQQPILKLAL 300

Qy        292 QRALQDASEVDEGGGQLPEEILELSGSRL 320
              |||||| :|||   |   :   |:| | |
Db        301 QRALQDIAEVDHSEGCFRQGTPEMSRSSL 329


RESULT 11
Q8R4G1
ID   Q8R4G1      PRELIMINARY;      PRT;   337 AA.
AC   Q8R4G1;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   G protein-coupled receptor SNSR.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY;
RX   MEDLINE=21853733; PubMed=11850634;
RA   Lembo P.M., Grazzini E., Groblewski T., O'Donnell D., Roy M.-O.,
RA   Zhang J., Hoffert C., Cao J., Schmidt R., Pelletier M., Labarre M.,
RA   Gosselin M., Fortin Y., Banville D., Shen S., Strom P., Payza K.,
RA   Dray A., Walker P., Ahmad S.;
RT   "Proenkephalin A gene products activate a new family of sensory
RT   neuron-specific GPCRs.";
RL   Nat. Neurosci. 5:201-209(2002).
DR   EMBL; AF474986; AAL86877.1; -.
KW   Receptor.
SQ   SEQUENCE   337 AA;  38702 MW;  B8D72EED92C65E2A CRC64;

  Query Match             51.0%;  Score 847;  DB 11;  Length 337;
  Best Local Similarity   53.4%;  Pred. No. 5.3e-69;
  Matches  172;  Conservative   48;  Mismatches   92;  Indels   10;  Gaps    5;

Qy          1 MDPTISTLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNA 60
              ||||||:| || | :| |    | :  |:|: |  |::|:|| || :|||||| |||| |
Db         15 MDPTISSLSTESTTLNKTGHPSC-RPILTLSFLVPIITLLGLAGNTIVLWLLGFRMRRKA 73

Qy         61 FSIYILNLAAADFLFLSGRLIYSLLSFISI----PHTISK-ILYPVMMFSYFAGLSFLSA 115
              |:|:|||: ||  ||    | ||:  ::      | :|| ||  |    | :||| |||
Db         74 ISVYVLNLSLADSFFLCCHFIDSLMRIMNFYGIYAHKLSKEILGNVAFIPYISGLSILSA 133

Qy        116 VSTERCLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQT 175
              :|||||||||||||| |||| ::||::|||:| || |  ||:|   |||       | :
Db        134 ISTERCLSVLWPIWYHCHRPRNMSAIICVLIWVLSFLMGILDWFFSGFLGETHHHLW-KN 192

Qy        176 SDFITVAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGIQ 235
              |||  |:|||| ::| |||| ||:|||||||: ||:|||||| |||:|:|:|||| |:
Db        193 VDFIVTAFLIFLFMLLFGSSLALLVRILCGSRRKPLSRLYVTISLTVMVYLICGLPLGLY 252

Qy        236 FFLFLWIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRAL 295
              ||  |  :     |||:: |:: || :|||||||||| ||||| |:  ::||:||:|||
Db        253 LFLLYWFGIHLHYPFCHIYQVTVLLSCVNSSANPIIYFLVGSFRHRKKHRSLKMVLKRAL 312
```

```
Qy        296 QDASEVDE---GGGQLPEEILE 314
              ::  | ||      | | ||| |
Db        313 EETPEEDEYTDSHVQKPTEISE 334


RESULT 12
Q91ZC2
ID   Q91ZC2      PRELIMINARY;      PRT;   338 AA.
AC   Q91ZC2;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   G protein-coupled receptor.
GN   MRGB2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6;
RX   MEDLINE=21435808; PubMed=11551509;
RA   Dong X., Han S.-K., Zylka M.J., Simon M.I., Anderson D.J.;
RT   "A Diverse Family of GPCRs Expressed in Specific Subsets of
RT   Nociceptive Somatosensory Neurons.";
RL   Cell 106:619-632(2001).
DR   EMBL; AY042200; AAK91796.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   338 AA;  38832 MW;  1B0A091D67C868B9 CRC64;

  Query Match             49.4%;  Score 821;  DB 11;  Length 338;
  Best Local Similarity   51.4%;  Pred. No. 1.2e-66;
  Matches  164;  Conservative   51;  Mismatches   92;  Indels   12;  Gaps    4;

Qy          6 STLDTELTPINGT---EETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNAFS 62
              |    | :| :||:   : ::|  :  :: :|: |:||||:  ||:||| || ||  |||:
Db         13 SAWKTNITVLNGSYYIDTSVCVTRNQAMILLSIIISLVGMGLNAIVLWFLGIRMHTNAFT 72

Qy         63 IYILNLAAADFLFLSGRLIYSLL----SFISIPHTISKILYPVMMFSYFAGLSFLSAVST 118
              :|||||| ||||:|  : :  ||      | ||    |  :|| | :|:| :||| || :|
Db         73 VYILNLAMADFLYLCSQFVICLLIAFYIFYSIDINIPLVLYVVPIFAYLSGLSILSTISI 132

Qy        119 ERCLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQTSDF 178
              ||||||:|||||||  || | ||: | :|| :|||  :||   || ||:  || ||:| |
Db        133 ERCLSVIWPIWYRCKRPRHTSAITCFVLWVMSLLLGLLEGKACGLLFNSFDSYWCETFDV 192

Qy        179 ITVAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGIQFFL 238
              ||  | :    ||||||| ||:|| |||::||:||||||| |||||||: |||||| : |
Db        193 ITNIWSVVFFGVLCGSSLTLLVRIFCGSQRIPMTRLYVTITLTVLVFLIFGLPFGIYWIL 252

Qy        239 FLWIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQ-NRQNLKLVLQRALQD 297
              : ||     |  |: :|  :||| :||  |||||| ||| | |:  |: |||:||||:||
Db        253 YQWISNFYYVEICNFYLEILFLSCVNSCMNPIIYFLVGSIRHRRFRRKTLKLLLQRAMQD 312

Qy        298 ASEVDEGG----GQLPEEI 312
              | :: |    : |||:
Db        313 TPEEEQSGNKSSSEHPEEL 331


RESULT 13
Q91ZC3
ID   Q91ZC3      PRELIMINARY;      PRT;   338 AA.
AC   Q91ZC3;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   G protein-coupled receptor.
GN   MRGB1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6;
RX   MEDLINE=21435808; PubMed=11551509;
RA   Dong X., Han S.-K., Zylka M.J., Simon M.I., Anderson D.J.;
RT   "A Diverse Family of GPCRs Expressed in Specific Subsets of
RT   Nociceptive Somatosensory Neurons.";
RL   Cell 106:619-632(2001).
DR   EMBL; AY042199; AAK91795.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   338 AA;  38750 MW;  9597037ED0BE8CE5 CRC64;

  Query Match             46.3%;  Score 769.5;  DB 11;  Length 338;
  Best Local Similarity   49.5%;  Pred. No. 5.9e-62;
  Matches  160;  Conservative   49;  Mismatches  103;  Indels   11;  Gaps    4;

Qy          1 MDPTISTLDTELTPINGTEE---TLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMR 57
              ||  |      :| :  : :   : |   : ::| |: |::|||| ||| ||| || :|
Db          1 MDLVIQDWTINITALKESNDNGISFCEVVSRTMTFLSLIIALVGLVGNATVLWFLGFQMS 60

Qy         58 RNAFSIYILNLAAADFLFLSGRLI---YSLLSFISIPHTISKILYPVMMFSYFAGLSFLS 114
              |||||:|||||| |||:|:  :::   | :|    |'          |:  :| :||| |:
Db         61 RNAFSVYILNLAGADFVFMCFQIVHCFYIILDIYFIPTNFFSSYTMVLNIAYLSGLSILT 120

Qy        115 AVSTERCLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQ 174
              :|||| |||:|||||| || | |||:| :|| |||: |:||   ||||:  :    |:
Db        121 VISTERFLSVMWPIWYRCQRPRHTSAVICTVLWVLSLVLSLLEGKECGFLYYTSGPGLCK 180

Qy        175 TSDFITVAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGI 234
              | | || |||| | ||| |||| |:: | ||  |:|:|||||||: ||||||: |||:||
Db        181 TFDLITTAWLIVLFVVLLGSSLALVLTIFCGLHKVPVTRLYVTIVFTVLVFLIFGLPYGI 240

Qy        235 QFFLFLWIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFR-QRQNRQNLKLVLQR 293
              :||  ||    :   |    |:|||| :|| ||||||| ||| |  | |: |||:|||
Db        241 YWFLLEWIREFHDNKPCGFRNVTIFLSCINSCANPIIYFLVGSIRHHRFQRKTLKLLLQR 300

Qy        294 ALQDASEVDE----GGGQLPEEI 312
              |:||: | :|    |  : | ||
Db        301 AMQDSPEEEECGEMGSSRRPREI 323


RESULT 14
Q91YB7
ID   Q91YB7      PRELIMINARY;      PRT;   331 AA.
AC   Q91YB7;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   G-protein coupled receptor.
GN   RC 56.1.3.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Bender E.;
RT   "Cloning and functional characterization of the rat orphan GPCR
RT   rc 56.1.3.";
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ311952; CAC84592.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Receptor.
```

```
SQ   SEQUENCE   331 AA;  37005 MW;  74BAEA8C557859AC CRC64;

  Query Match           45.8%;  Score 761.5;  DB 11;  Length 331;
  Best Local Similarity  51.0%;  Pred. No. 3.1e-61;
  Matches  172;  Conservative   39;  Mismatches   91;  Indels   35;  Gaps    7;

Qy          1 MDPTISTLD-TELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRN 59
              :: : ||:  |   |:: |       :||   :|  |  |||| |||:| |||| |: ||
Db         13 LNTSASTIAVTTTNPMDKTIPGSFNGRTLIPNLLIIISGLVGLIGNAMVFWLLGFRLARN 72

Qy         60 AFSIYILNLAAADFLFLSGRLIYSLLSFISIPHTISKILYP----------VMMFSYFAG 109
              |||:|||||| ||||||   :| | |  :       |  ||          |||  | ||
Db         73 AFSVYILNLALADFLFLLCHIIDSTLLLL-------KFSYPNIIFLPCFNTVMMVPYIAG 125

Qy        110 LSFLSAVSTERCLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGAD 169
              || |||:||||||||: |||||| || | | |:|  :| ||||  ||   ||||    |
Db        126 LSMLSAISTERCLSVVCPIWYRCRRPKHTSTVMCSAIWVLSLLICILNRYFCGFL----D 181

Qy        170 SAW-----CQTSDFITVAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLV 224
              : :      |  |:| | | |||| |||| ||| ||:|  ||: :: ||||| ||:|||||
Db        182 TKYEKDNRCLASNFFTAACLIFLFVVLCLSSLALLVRSFCGAGRMKLTRLYATIMLTVLV 241

Qy        225 FLLCGLPFGIQFFLFLWIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNR 284
              |||||||||| :|| :|| :|      ::| :: |:|:|| ||||||||||||| |
Db        242 FLLCGLPFGIHWFLLIWIKIDYGKFAYGLYLAALVLTAVNSCANPIIYFFVGSFR-HQKH 300

Qy        285 QNLKLVLQRALQDASEVDEGGGQLPEEILELSGSRLE 321
              | ||:|||||||  |.        |  :|:| |::|
Db        301 QTLKMVLQRALQDTPET-------AENTVEMSSSKVE 330


RESULT 15
Q91WW5
ID   Q91WW5      PRELIMINARY;      PRT;   304 AA.
AC   Q91WW5;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   RF-amide G protein-coupled receptor.
GN   MRGA1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6; TISSUE=DORSAL ROOT GANGLION;
RX   MEDLINE=21435808; PubMed=11551509;
RA   Dong X., Han S., Zylka M.J., Simon M.I., Anderson D.J.;
RT   "A diverse family of gpcrs expressed in specific subsets of
RT   nociceptive sensory neurons.";
RL   Cell 106:619-632(2001).
DR   EMBL; AY042191; AAK91787.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   304 AA;  34381 MW;  C56CBF879067A52B CRC64;

  Query Match           45.5%;  Score 756.5;  DB 11;  Length 304;
  Best Local Similarity  53.0%;  Pred. No. 8e-61;
  Matches  160;  Conservative   43;  Mismatches   84;  Indels   15;  Gaps    4;

Qy         28 LSLTVL-----TCIVSLVGLTGNAVVLWLLGCRMRRNAFSIYILNLAAADFLFLSGRLIY 82
              :::|:|       |  ||||||| :| ||||  : |||||:|||||| ||| || | :|
Db          9 INITILIPNLMIIIFGLVGLTGNGIVFWLLGFCLHRNAFSVYILNLALADFFFLLGHIID 68

Qy         83 SLLSFISIPHTISKIL--YPVMMFSYFAGLSFLSAVSTERCLSVLWPIWYRCHRPTHLSA 140
              |:|  ::: : |: :|  | :||  | |||| |||:||||||||| |||| |||| | |
Db         69 SILLLLNVFYPITFLLCFYTIMMVLYIAGLSMLSAISTERCLSVLCPIWYHCHRPEHTST 128

Qy        141 VVCVLLWALSLLRSILEWMLCGFLFSG-ADSAWCQTSDFITVAWLIFLCVVLCGSSLVLL 199
              |:| ::| |||| ||    |||| :   :   |   :| | |:|:|| |||| ||| |:
Db        129 VMCAVIWVLSLLICILNSYFCGFLNTQYKNENGCLALNFFTAAYLMFLFVVLCLSSLALV 188

Qy        200 IRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGIQFFLFLWIHVDREVLFCHVHLVSIF 259
              |: ||: :| ||||||||:|::|||||||||||| :||   |  |  |     :| |:
Db        189 ARLFCGTGQIKLTRLYVTIILSILVFLLCGLPFGIHWFLLFKIKDDFHVFDLGFYLASVV 248

Qy        260 LSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASEVDEGGGQLPEEILELSGSR 319
              |:|:|| |||||||||||||| |   | ||:||| ||||  |         : ::|:| |:
Db        249 LTAINSCANPIIYFFVGSFRHRLKHQTLKMVLQNALQDTPET-------AKIMVEMSRSK 301

Qy        320 LE 321
              ||
Db        302 SE 303


Search completed: July  2, 2003, 19:22:29
Job time : 82 secs
```

GenCore version 5.1.6


OM protein - protein search, using sw model

Run on: July 2, 2003, 19:14:05 ; Search time 11 Seconds
(without alignments)
1214.125 Million cell updates/sec

Title: US-09-849-869A-16
Perfect score: 1661
Sequence: 1 MDPTISTLDTELTPINGTEE.....EGGGQLPEEILELSGSRLEQ 322

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 484.5 | 29.2 | 324 | 1 | MAS_RAT | P12526 rattus norv |
| 2 | 462 | 27.8 | 378 | 1 | MRG_HUMAN | P35410 homo sapien |
| 3 | 461.5 | 27.8 | 325 | 1 | MAS_HUMAN | P04201 homo sapien |
| 4 | 456.5 | 27.5 | 324 | 1 | MAS_MOUSE | P30554 mus musculu |
| 5 | 418 | 25.2 | 343 | 1 | RTA_RAT | P23749 rattus norv |
| 6 | 279.5 | 16.8 | 395 | 1 | GP44_HUMAN | Q9y5y4 homo sapien |
| 7 | 262 | 15.8 | 382 | 1 | GP44_MOUSE | Q9z2j6 mus musculu |
| 8 | 256.5 | 15.4 | 373 | 1 | CML1_HUMAN | Q99788 homo sapien |
| 9 | 256 | 15.4 | 349 | 1 | FML2_PANTR | P79243 pan troglod |
| 10 | 252 | 15.2 | 340 | 1 | C5AR_GORGO | P79175 gorilla gor |
| 11 | 252 | 15.2 | 349 | 1 | FML2_GORGO | P79178 gorilla gor |
| 12 | 252 | 15.2 | 349 | 1 | FML2_PONPY | P79237 pongo pygma |
| 13 | 252 | 15.2 | 473 | 1 | C3AR_RAT | O55197 rattus norv |
| 14 | 251 | 15.1 | 353 | 1 | FML2_HUMAN | P25039 homo sapien |
| 15 | 250 | 15.1 | 340 | 1 | C5AR_MACMU | P79138 macaca mula |
| 16 | 250 | 15.1 | 348 | 1 | FML1_PONPY | P79236 pongo pygma |
| 17 | 250 | 15.1 | 350 | 1 | C5AR_HUMAN | P21730 homo sapien |
| 18 | 249.5 | 15.0 | 340 | 1 | C5AR_PONPY | P79234 pongo pygma |
| 19 | 248 | 14.9 | 348 | 1 | FML1_GORGO | P79177 gorilla gor |
| 20 | 248 | 14.9 | 351 | 1 | FML1_HUMAN | P25090 homo sapien |
| 21 | 247.5 | 14.9 | 345 | 1 | C5AR_CAVPO | O70129 cavia porce |
| 22 | 244 | 14.7 | 348 | 1 | FML1_PANTR | P79242 pan troglod |
| 23 | 243.5 | 14.7 | 369 | 1 | SSR2_RAT | P30680 rattus norv |
| 24 | 243 | 14.6 | 340 | 1 | C5AR_PANTR | P79240 pan troglod |
| 25 | 242.5 | 14.6 | 350 | 1 | C5AR_RABIT | Q9tue1 oryctolagus |
| 26 | 242.5 | 14.6 | 355 | 1 | GPR1_HUMAN | P46091 homo sapien |
| 27 | 241.5 | 14.5 | 369 | 1 | SSR2_MOUSE | P30875 mus musculu |
| 28 | 241 | 14.5 | 349 | 1 | FML2_MACMU | P79191 macaca mula |
| 29 | 241 | 14.5 | 475 | 1 | C3AR_CAVPO | O88680 cavia porce |
| 30 | 240.5 | 14.5 | 355 | 1 | GPR1_MACMU | O97664 macaca mula |
| 31 | 238 | 14.3 | 351 | 1 | FML1_MOUSE | O08790 mus musculu |
| 32 | 237 | 14.3 | 348 | 1 | FML1_MACMU | P79190 macaca mula |
| 33 | 236 | 14.2 | 364 | 1 | FMLR_MOUSE | P33766 mus musculu |
| 34 | 236 | 14.2 | 381 | 1 | GP34_HUMAN | Q9upc5 homo sapien |
| 35 | 233 | 14.0 | 356 | 1 | GP32_HUMAN | O75388 homo sapien |
| 36 | 232 | 14.0 | 375 | 1 | GP34_MOUSE | Q9r1k6 mus musculu |
| 37 | 231.5 | 13.9 | 346 | 1 | FMLR_MACMU | P79189 macaca mula |
| 38 | 230 | 13.8 | 352 | 1 | C5AR_RAT | P97520 rattus norv |
| 39 | 230 | 13.8 | 371 | 1 | CML1_MOUSE | P97468 mus musculu |
| 40 | 229 | 13.8 | 346 | 1 | FMLR_PONPY | P79235 pongo pygma |
| 41 | 229 | 13.8 | 352 | 1 | C5AR_CANFA | P30992 canis famil |
| 42 | 226 | 13.6 | 369 | 1 | SSR2_PIG | P34994 sus scrofa |
| 43 | 225 | 13.5 | 352 | 1 | FMLR_RABIT | Q05394 oryctolagus |
| 44 | 224.5 | 13.5 | 355 | 1 | IL8A_RABIT | P21109 oryctolagus |
| 45 | 224 | 13.5 | 362 | 1 | SSR5_MOUSE | O08858 mus musculu |

ALIGNMENTS

RESULT 1
MAS_RAT
ID MAS_RAT STANDARD; PRT; 324 AA.
AC P12526;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE MAS proto-oncogene.
GN MAS1 OR MAS-1 OR MAS.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88276953; PubMed=2455902;
RA Young D., O'Neill K., Jessell T., Wigler M.;
RT "Characterization of the rat mas oncogene and its high-level
RT expression in the hippocampus and cerebral cortex of rat brain.";
RL Proc. Natl. Acad. Sci. U.S.A. 85:5339-5342(1988).
CC -!- FUNCTION: NOT KNOWN.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -!- CAUTION: WAS ORIGINALLY THOUGHT TO BE A RECEPTOR FOR ANGIOTENSIN
CC II.
CC ---

CC ---
DR EMBL; J03823; AAA41573.1; -.
DR PIR; A31816; TVRTAS.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Proto-oncogene.

| | | | | |
|---|---|---|---|---|
| FT | DOMAIN | 1 | 35 | EXTRACELLULAR (POTENTIAL). |
| FT | TRANSMEM | 36 | 60 | 1 (POTENTIAL). |
| FT | DOMAIN | 61 | 64 | CYTOPLASMIC (POTENTIAL). |
| FT | TRANSMEM | 65 | 86 | 2 (POTENTIAL). |
| FT | DOMAIN | 87 | 103 | EXTRACELLULAR (POTENTIAL). |
| FT | TRANSMEM | 104 | 127 | 3 (POTENTIAL). |
| FT | DOMAIN | 128 | 148 | CYTOPLASMIC (POTENTIAL). |
| FT | TRANSMEM | 149 | 171 | 4 (POTENTIAL). |
| FT | DOMAIN | 172 | 184 | EXTRACELLULAR (POTENTIAL). |
| FT | TRANSMEM | 185 | 205 | 5 (POTENTIAL). |
| FT | DOMAIN | 206 | 223 | CYTOPLASMIC (POTENTIAL). |
| FT | TRANSMEM | 224 | 244 | 6 (POTENTIAL). |
| FT | DOMAIN | 245 | 262 | EXTRACELLULAR (POTENTIAL). |
| FT | TRANSMEM | 263 | 283 | 7 (POTENTIAL). |

```
FT   DOMAIN      284    324       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD      5      5       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     16     16       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     21     21       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   324 AA;  37130 MW;  AD810229BF1E7D36 CRC64;

  Query Match             29.2%;  Score 484.5;  DB 1;  Length 324;
  Best Local Similarity   39.2%;  Pred. No. 1.9e-25;
  Matches  112;  Conservative   61;  Mismatches   86;  Indels   27;  Gaps   10;

Qy         37 VSLVGLTGNAVVLWLLGCRMRRNAFSIYILNLAAADFLFLSGRLI----YSLLSFISIPH 92
              :| :|   | ::|| |  ||||| |::|| :|: ||   |     |     |:|   :|  |
Db         41 ISPLGFVENGILLWFLCFRMRRNPFTVYITHLSIADISLLFCIFILSIDYALDYELSSGH 100

Qy         93 TISKILYPV-MMFSYFAGLSFLSAVSTERCLSVLWPIWYRCHRPTHLSAVVCVLLWALSL 151
                :  :   |  :| |  ||  |:|:| |||||||:||||||||| | || || ||||||
Db        101 YYTIVTLSVTFLFGYNTGLYLLTAISVERCLSVLYPIWYRCHRPKHQSAFVCALLWALSC 160

Qy        152 LRSILEWMLCGFLFSGADSAWCQTSD------FITV-AWLIFLCVVLCGSSLVLLIRILC 204
              | : :|:::|  : || :|     ||       || : ::|:|  ::|  || :|:::|
Db        161 LVTTMEYVMC--IDSGEESH--SQSDCRAVIIFIAILSFLVFTPLMLV-SSTILVVKIRK 215

Qy        205 GSRKIPLTRLYVTILLTVLVFLLCGLPFGIQFFLF--LWIHVDREVLFCHVHLVSIFLSA 262
                :    ::||: |::|:::||:  :|  : : |:    |         | ::| :|:  |
Db        216 NTWASHSSKLYIVIMVTIIIFLIFAMPMRVLYLLYYEYW------STFGNLHNISLLFST 269

Qy        263 LNSSANPIIYFFVGSSFRQRQNRQNLKLVLQRALQDASEV--DEGGG 306
              :|||||| ||||||| :::: |::||:|| || :|  :    || |
Db        270 INSSANPFIYFFVGSSKKKRFRESLKVVLTRAFKDEMQPRRQEGNG 315


RESULT 2
MRG_HUMAN
ID   MRG_HUMAN      STANDARD;      PRT;   378 AA.
AC   P35410;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   MAS-related G protein-coupled receptor MRG.
GN   MRG.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92130997; PubMed=1723144;
RA   Monnot C., Weber V., Stinnakre J., Bihoreau C., Teutsch B., Corvol P.,
RA   Clauser E.;
RT   "Cloning and functional characterization of a novel mas-related gene,
RT   modulating intracellular angiotensin II actions.";
RL   Mol. Endocrinol. 5:1477-1487(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Whitaker H.;
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC       MOST SIMILAR TO MAS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; S78653; AAB21255.1; -.
DR   EMBL; AL035542; CAB44503.1; -.
DR   PIR; A39485; A39485.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein.
FT   DOMAIN        1     77       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     78    101       1 (POTENTIAL).
FT   DOMAIN      102    109       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    110    136       2 (POTENTIAL).
FT   DOMAIN      137    154       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    155    169       3 (POTENTIAL).
FT   DOMAIN      170    191       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    192    207       4 (POTENTIAL).
FT   DOMAIN      208    221       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    222    248       5 (POTENTIAL).
FT   DOMAIN      249    264       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    265    286       6 (POTENTIAL).
FT   DOMAIN      287    297       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    298    317       7 (POTENTIAL).
FT   DOMAIN      318    378       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD     54     54       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     57     57       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   378 AA;  42410 MW;  E08A606294B61474 CRC64;

  Query Match             27.8%;  Score 462;  DB 1;  Length 378;
  Best Local Similarity   37.4%;  Pred. No. 6.6e-24;
  Matches  114;  Conservative   56;  Mismatches  101;  Indels   34;  Gaps    9;

Qy         16 NGTEETL-------CYKQTLSLTVL---TCIVSLVGLTGNAVVLWLLGCRMRRNAFSIYI 65
              | | ||:       :| | | ::     :||| |:  |  | ||| |    | : :||
Db         54 NETNETIHMQMSMAVGQQALPLNIIAPKAVLVSLCGVLLNGTVFWLLCCG-ATNPYMVYI 112

Qy         66 LNLAAADFLFLS----GRLIYSLLSFISIPHTISKILYPVMMFSYFAGLSFLSAVSTERC 121
              |:| ||| ::|     | |  :||::  :    |   |  :  ||:   |  | |:|||||
Db        113 LHLVAADVIYLCCSAVGFLQVTLLTYHGVVFFIPDFLAILSPFSFEVCLCLLVAISTERC 172

Qy        122 LSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQTSD---F 178
              : ||:||||||||| : | ||| |:|  |    :|:: :  :          |      |
Db        173 VCVLFPIWYRCHRPKYTSNVVCTLIWGLPFCINIVKSLFLTY--------WKHVKACVIF 224

Qy        179 ITVAWLI--FLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGIQF 236
              : :: |     | :|:| ||| |||| || |::   ||:|  : ::  :|||  ||  :
Db        225 LKLSGLFHAILSLVMCVSSLTLLIRFLCCSQQQKATRVYAVVQISAPMFLLWALPLSVAP 284

Qy        237 FLFLWIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQ 296
              :     | ::     :|:|:||  :|||||||||||||| |::: :::|:::|||||
Db        285 LI-----TDFKMFVTTSYLISLFL-IINSSANPIIYFFVGSLRKKRLKESLRVILQRALA 338

Qy        297 DASEV 301
              |  ||
Db        339 DKPEV 343


RESULT 3
MAS_HUMAN
ID   MAS_HUMAN      STANDARD;      PRT;   325 AA.
AC   P04201;
DT   20-MAR-1987 (Rel. 04, Created)
DT   20-MAR-1987 (Rel. 04, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   MAS proto-oncogene.
GN   MAS1 OR MAS.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86218084; PubMed=3708691;
RA   Young D., Waitches G., Birchmeier C., Fasano O., Wigler M.;
RT   "Isolation and characterization of a new cellular oncogene encoding a
RT   protein with multiple potential transmembrane domains.";
RL   Cell 45:711-719(1986).
RN   [2]
RP   POSSIBLE FUNCTION.
RX   MEDLINE=88334724; PubMed=3419518;
```

```
RA   Jackson T.R., Blair L.A.C., Marshall J., Goedert M., Hanley M.R.;
RT   "The mas oncogene encodes an angiotensin receptor.";
RL   Nature 335:437-440(1988).
CC   -!- FUNCTION: NOT KNOWN.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- DISEASE: THE MAS ONCOGENE HAS A WEAK FOCUS-INDUCING ACTIVITY IN
CC       TRANSFECTED NIH 3T3 CELLS.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   -!- CAUTION: Was originally (Ref.1) thought to be a receptor for
CC       angiotensin II.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M13150; AAA36199.1; -.
DR   PIR; A01375; TVHUAS.
DR   Genew; HGNC:6899; MAS1.
DR   MIM; 165180; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Proto-oncogene.
FT   DOMAIN        1     36       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     37     61       1 (POTENTIAL).
FT   DOMAIN       62     65       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     66     86       2 (POTENTIAL).
FT   DOMAIN       87    104       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    105    128       3 (POTENTIAL).
FT   DOMAIN      129    149       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    150    172       4 (POTENTIAL).
FT   DOMAIN      173    185       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    186    206       5 (POTENTIAL).
FT   DOMAIN      207    224       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    225    245       6 (POTENTIAL).
FT   DOMAIN      246    263       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    264    284       7 (POTENTIAL).
FT   DOMAIN      285    325       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD      5      5       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     16     16       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     22     22       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   325 AA;  37465 MW;  3368E7B174744B83 CRC64;

  Query Match           27.8%;  Score 461.5;  DB 1;  Length 325;
  Best Local Similarity  38.5%;  Pred. No. 6.2e-24;
  Matches  104;  Conservative   59;  Mismatches   92;  Indels   15;  Gaps    5;

Qy         37 VSLVGLTGNAVVLWLLGCRMRRNAFSIYILNLAAADFLFLSGRLI----YSLLSFISIPH 92
              :| ||    | ::|| |  ||||| |::|| :|: ||    |     |    |:|   :|  |
Db         42 ISPVGFVENGILLWFLCFRMRRNPFTVYITHLSIADISLLFCIFILSIDYALDYELSSGH 101

Qy         93 TISKILYPV-MMFSYFAGLSFLSAVSTERCLSVLWPIWYRCHRPTHLSAVVCVLLWALSL 151
              : :   |  :| |  ||  |:|:| |||||||:||||||||| : ||:|| |||||
Db        102 YYTIVTLSVTFLFGYNTGLYLLTAISVERCLSVLYPIWYRCHRPKYQSALVCALLWALSC 161

Qy        152 LRSILEWMLCGFLFSGADSAWCQTSDFITVAWLIFLCV--VLCGSSLVLLIRILCGSRKI 209
              | : :|:::|      : |     :  | :| | ||    ::  || :|:::|   :
Db        162 LVTTMEYVMCIDREEESHSRNDCRAVIIFIAILSFLVFTPLMLVSSTILVVKIRKNTWAS 221

Qy        210 PLTRLYVTILLTVLVFLLCGLPFGIQFFLF--LWIHVDREVLFCHVHLVSIFLSALNSSA 267
              ::||: |::|:::||:  :|  : : |:    |        | ::| :|:  | :||||
Db        222 HSSKLYIVIMVTIIIFLIFAMPMRLLYLLYYEYW------STFGNLHHISLLFSTINSSA 275

Qy        268 NPIIYFFVGSFRQRQNRQNLKLVLQRALQD 297
              || ||||||| :::: :::||:|| || :|
Db        276 NPFIYFFVGSSKKKRFKESLKVVLTRAFKD 305


RESULT 4
MAS_MOUSE
ID   MAS_MOUSE      STANDARD;      PRT;   324 AA.
AC   P30554; O35944;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   MAS proto-oncogene.
GN   MAS1 OR MAS-1 OR MAS.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/c; TISSUE=Liver;
RX   MEDLINE=95094925; PubMed=8001672;
RA   Metzger R., Bader M., Ludwig T., Berberich C., Bunnemann B.,
RA   Ganten D.;
RT   "Expression of the mouse and rat mas proto-oncogene in the brain and
RT   peripheral tissues.";
RL   FEBS Lett. 357:27-32(1995).
RN   [2]
RP   SEQUENCE OF 1-25 FROM N.A.
RC   STRAIN=BCBA; TISSUE=Testis;
RX   MEDLINE=97422605; PubMed=9268631;
RA   Schweifer N., Valk P.J., Delwel R., Cox R., Francis F.,
RA   Meier-Ewert S., Lehrach H., Barlow D.P.;
RT   "Characterization of the C3 YAC contig from proximal mouse chromosome
RT   17 and analysis of allelic expression of genes flanking the imprinted
RT   Igf2r gene.";
RL   Genomics 43:285-297(1997).
CC   -!- FUNCTION: NOT KNOWN.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   -!- CAUTION: WAS ORIGINALLY THOUGHT TO BE A RECEPTOR FOR ANGIOTENSIN
CC       II.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X67735; CAA47964.1; -.
DR   EMBL; U96273; AAB69120.1; -.
DR   PIR; S29619; S29619.
DR   MGD; MGI:96918; Mas1.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Proto-oncogene.
FT   DOMAIN        1     35       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     36     60       1 (POTENTIAL).
FT   DOMAIN       61     64       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     65     86       2 (POTENTIAL).
FT   DOMAIN       87    103       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    104    127       3 (POTENTIAL).
FT   DOMAIN      128    148       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    149    171       4 (POTENTIAL).
FT   DOMAIN      172    184       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    185    205       5 (POTENTIAL).
FT   DOMAIN      206    223       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    224    244       6 (POTENTIAL).
FT   DOMAIN      245    262       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    263    283       7 (POTENTIAL).
FT   DOMAIN      284    324       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD      5      5       N-LINKED (GLCNAC. . .) (POTENTIAL).
```

```
FT   CARBOHYD      16      16       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      21      21       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT      19      19       S -> I (IN REF. 2).
SQ   SEQUENCE   324 AA;  36904 MW;  24F4AB7299E6016F CRC64;

  Query Match             27.5%;  Score 456.5;  DB 1;  Length 324;
  Best Local Similarity   38.1%;  Pred. No. 1.3e-23;
  Matches  109;  Conservative   61;  Mismatches   89;  Indels   27;  Gaps   10;

Qy         37 VSLVGLTGNAVVLWLLGCRMRRNAFSIYILNLAAADFLFLSGRLI----YSLLSFISIPH 92
              :| :|   | ::|| |  ||||| |::|| :|: ||   |     |    |:|   :|  |
Db         41 ISPLGFVENGILLWFLCFRMRRNPFTVYITHLSMADISLLFCIFILSIDYALDYELSSGH 100

Qy         93 TISKILYPV-MMFSYFAGLSFLSAVSTERCLSVLWPIWYRCHRPTHLSAVVCVLLWALSL 151
                : :   |  :| |  ||  |:|:| |||||||:|||  ||| | || || || |||
Db        101 HYTIVTLSVTFLFGYNTGLYLLTAISVERCLSVLYPIWYTSHRPKHQSAFVCALLCALSC 160

Qy        152 LRSILEWMLCGFLFSGADSAWCQTSD------FITV-AWLIFLCVVLCGSSLVLLIRILC 204
              | : :|:::|  : || :|     ||       || : ::|:|  ::|  || :|:::|
Db        161 LVTTMEYVMC--IDSGEESH--SRSDCRAVIIFIAILSFLVFTPLMLVSSS-ILVVKIRK 215

Qy        205 GSRKIPLTRLYVTILLTVLVFLLCGLPFGIQFFLF--LWIHVDREVLFCHVHLVSIFLSA 262
                :    ::||: |::|:::||:  :|  : : |:   |              | ::| :|:  |
Db        216 NTWASHSSKLYIVIMVTIIIFLIFAMPMRVLYLLYYEYW------SAFGNLHNISLLFST 269

Qy        263 LNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASEV--DEGGG 306
              :|||||| ||||||| :::: |::||:|| || :|  :    || |
Db        270 INSSANPFIYFFVGSSKKKRFRESLKVVLTRAFKDEMQPRRQEGNG 315


RESULT 5
RTA_RAT
ID   RTA_RAT        STANDARD;      PRT;   343 AA.
AC   P23749;
DT   01-NOV-1991 (Rel. 20, Created)
DT   01-NOV-1991 (Rel. 20, Last sequence update)
DT   01-JUL-1993 (Rel. 26, Last annotation update)
DE   Probable G protein-coupled receptor RTA.
GN   RTA.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Aorta;
RX   MEDLINE=90222168; PubMed=2109324;
RA   Ross P.C., Figler R.A., Corjay M.H., Barber C.M., Adam N.,
RA   Harcus D.R., Lynch K.R.;
RT   "RTA, a candidate G protein-coupled receptor: cloning, sequencing,
RT   and tissue distribution.";
RL   Proc. Natl. Acad. Sci. U.S.A. 87:3052-3056(1990).
CC   -!- FUNCTION: ORPHAN RECEPTOR.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: GUT, VAS DEFERENS, UTERUS, AND AORTA BUT ONLY
CC       BARELY DETECTABLE IN LIVER, KIDNEY, LUNG, AND SALIVARY GLAND. IN
CC       THE BRAIN, RTA IS MARKEDLY ABUNDANT IN THE CEREBELLUM.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC       MOST SIMILAR TO MAS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M35297; AAA42087.1; -.
DR   EMBL; M35298; AAA42088.1; -.
DR   PIR; A35639; A35639.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein.
FT   DOMAIN        1      44       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     45      66       1 (POTENTIAL).
FT   DOMAIN       67      82       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     83     104       2 (POTENTIAL).
FT   DOMAIN      105     123       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    124     144       3 (POTENTIAL).
FT   DOMAIN      145     160       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    161     181       4 (POTENTIAL).
FT   DOMAIN      182     198       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    199     220       5 (POTENTIAL).
FT   DOMAIN      221     241       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    242     263       6 (POTENTIAL).
FT   DOMAIN      264     273       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    274     294       7 (POTENTIAL).
FT   DOMAIN      295     343       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD      4       4       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   343 AA;  38364 MW;  E4630007770941F4 CRC64;

  Query Match             25.2%;  Score 418;  DB 1;  Length 343;
  Best Local Similarity   35.0%;  Pred. No. 4.7e-21;
  Matches   99;  Conservative   57;  Mismatches  117;  Indels   10;  Gaps    6;

Qy         36 IVSLVGLTGNAVVLWLLGCRMRRNAFSIYILNLAAADFLFLSGRLIYSLL---SFI-SIP 91
              :: | || || :|||  |  ::|  |||| |:||:|| ::|  : : :||   :|: | |
Db         53 LLCLCGLVGNGLVLWFFGFSIKRTPFSIYFLHLASADGIYLFSKAVIALLNMGTPLGSFP 112

Qy         92 HTISKILYPVMMFSYFAGLSFLSAVSTERCLSVLWPIWYRCHRPTHLSAVVCVLLWALSL 151
                : ::   | : ::|||:| | |:| |||:||::|:||   ||  ||| || ||| ||
Db        113 DYVRRVSRIVGLCTFFAGVSLLPAISIERCVSVIFPMWYWRRRPKRLSAGVCALLWLLSF 172

Qy        152 LRSILEWMLCGFLFSGADSAWCQTSDFITVAWLIFL-CVVLCGSSLVLLIRILCGSRKIP 210
              | : :   | ||   |   |   |      | || | ::   | |:: : | :|:
Db        173 LVTSIHNYFCMFLGHEASGTACLNMDISLGILLFFLFCPLMVLPCLALILHVECRARRRQ 232

Qy        211 LT-RLYVTILLTVLVFLLCGLPFGIQFFLFLWIHVDREVLFCHVHLVSIFLSALNSSANP 269
                : :|   :|  | |||:  :  || :||| |:    ::       |:     :|||| |
Db        233 RSAKLNHVVLAIVSVFLVSSIYLGIDWFLF-WVF---QIPAPFPEYVTDLCICINSSAKP 288

Qy        270 IIYFFVGSFRQRQNRQNLKLVLQRALQDASEVDEGGGQLPEEI 312
              |:||  |  : ::  : |::| ||||:| :|  :      |  :
Db        289 IVYFLAGRDKSQRLWEPLRVVFQRALRDGAEPGDAASSTPNTV 331


RESULT 6
GP44_HUMAN
ID   GP44_HUMAN     STANDARD;      PRT;   395 AA.
AC   Q9Y5Y4; O94765;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Putative G protein-coupled receptor GPR44 (Chemoattractant receptor-
DE   homologous molecule expressed on Th2 cells).
GN   GPR44 OR CRTH2 OR DL1R.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99156852; PubMed=10036181;
RA   Marchese A., Sawzdargo M., Nguyen T., Cheng R., Heng H.H., Nowak T.,
RA   Im D.S., Lynch K.R., George S.R., O'dowd B.F.;
RT   "Discovery of three novel orphan G-protein-coupled receptors.";
RL   Genomics 56:12-21(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Blood;
RX   MEDLINE=99138803; PubMed=9973380;
RA   Nagata K., Tanaka K., Ogawa K., Kemmotsu K., Imai T., Yoshie O.,
RA   Abe H., Tada K., Nakamura M., Sugamura K., Takano S.;
```

```
RT   "Selective expression of a novel surface molecule by human Th2 cells
RT   in vivo.";
RL   J. Immunol. 162:1278-1286(1999).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RA   Methner A., Schroeder S.;
RT   "Tissue expression and chromosomal organization of a novel G protein-
RT   coupled receptor.";
RL   Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: ORPHAN RECEPTOR.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF118265; AAD21055.1; ALT_INIT.
DR   EMBL; AB008535; BAA74518.1; -.
DR   EMBL; AF144308; AAD34539.1; -.
DR   Genew; HGNC:4502; GPR44.
DR   MIM; 604837; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein.
FT   DOMAIN        1     33       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     34     56       1 (POTENTIAL).
FT   DOMAIN       57     67       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     68     89       2 (POTENTIAL).
FT   DOMAIN       90    106       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    107    127       3 (POTENTIAL).
FT   DOMAIN      128    146       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    147    168       4 (POTENTIAL).
FT   DOMAIN      169    210       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    211    231       5 (POTENTIAL).
FT   DOMAIN      232    247       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    248    269       6 (POTENTIAL).
FT   DOMAIN      270    288       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    289    308       7 (POTENTIAL).
FT   DOMAIN      309    395       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD      4      4       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     25     25       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DISULFID    104    182       POTENTIAL.
FT   CONFLICT    375    395       GSCAASPQTGPLNRALSSTSS -> VQLRSVPADGPPEPGA
FT                                EQHLELEPGPRRAALTRESITRVPRFNSISGLLPQ (IN
FT                                REF. 1).
SQ   SEQUENCE   395 AA;  43239 MW;  9DBBB53B2008C1D1 CRC64;

  Query Match             16.8%;  Score 279.5;  DB 1;  Length 395;
  Best Local Similarity   27.7%;  Pred. No. 6.8e-12;
  Matches   84;  Conservative   66;  Mismatches  114;  Indels   39;  Gaps   11;

Qy         36 IVSLVGLTGNAVVLWLLGCRMRRNAFSIYILNLAAADFLFLSGRLIYSLLSFISIPH--- 92
              : ||:||  | |:|:::|||||:   : ::|:|| :| |  :    ::   |::: |
Db         41 LASLLGLVENGVILFVVGCRMRQTVVTTWVLHLALSDLLASASLPFFTY--FLAVGHSWE 98

Qy         93 ---TISKILYPVMMFSYFAGLSFLSAVSTERCLSVLWPIWYRCHRPTHLSAVVCVLLWAL 149
                 |  |:   :   : ||    |||:| :||| |: |:| : ||     :  ||::||||
Db         99 LGTTFCKLHSSIFFLNMFASGFLLSAISLDRCLQVVRPVWAQNHRTVAAAHKVCLVLWAL 158

Qy        150 SLLRSILEW-----------MLCGF----LFSGAD-SAWCQT-------SDFITVAWLI 185
              ::| ::  :           ::| :    |  | |  | :       | |: :|:|:
Db        159 AVLNTVPYFVFRDTISRLDGRIMCYYNVLLLNPGPDRDATCNSRQAALAVSKFL-LAFLV 217

Qy        186 FLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGIQFFLFLWIHVD 245
```

```
                | ::  |   : :|:    |: |  |    :   |  | ||  |: :   |    | :
Db        218 PLAII-ASSHAAVSLRLQHRGRRRP-GRFVRLVAAVVAAFALCWGPYHVFSLLEARAHAN 275

Qy        246 ---REVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASEVD 302
                | :::  :  |:  |:  || |||::|       |: |::|: ||:  | | ||:
Db        276 PGLRPLVWRGLPFVT-SLAFFNSVANPVLYVLTCPDMLRKLRRSLRTVLESVLVDDSELG 334

Qy        303 EGG 305
                |
Db        335 GAG 337


RESULT 7
GP44_MOUSE
ID   GP44_MOUSE     STANDARD;      PRT;   382 AA.
AC   Q9Z2J6;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Putative G protein-coupled receptor GPR44 (Chemoattractant receptor-
DE   homologous molecule expressed on Th2 cells).
GN   GPR44 OR CRTH2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SvJ;
RX   MEDLINE=99132318; PubMed=9931443;
RA   Abe H., Takeshita T., Nagata K., Arita T., Endo Y., Fujita T.,
RA   Takayama H., Kubo M., Sugamura K.;
RT   "Molecular cloning, chromosome mapping and characterization of the
RT   mouse CRTH2 gene, a putative member of the leukocyte chemoattractant
RT   receptor family.";
RL   Gene 227:71-77(1999).
CC   -!- FUNCTION: ORPHAN RECEPTOR.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF054507; AAD13525.1; -.
DR   MGD; MGI:1330275; Gpr44.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein.
FT   DOMAIN        1     32       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     33     55       1 (POTENTIAL).
FT   DOMAIN       56     66       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     67     88       2 (POTENTIAL).
FT   DOMAIN       89    105       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    106    126       3 (POTENTIAL).
FT   DOMAIN      127    145       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    146    167       4 (POTENTIAL).
FT   DOMAIN      168    209       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    210    230       5 (POTENTIAL).
FT   DOMAIN      231    246       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    247    268       6 (POTENTIAL).
FT   DOMAIN      269    287       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    288    307       7 (POTENTIAL).
FT   DOMAIN      308    357       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD      3      3       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     21     21       N-LINKED (GLCNAC. . .) (POTENTIAL).
```

```
FT   CARBOHYD     24     24       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DISULFID    103    181       POTENTIAL.
SQ   SEQUENCE   382 AA;  42949 MW;  8CCB898B93795B04 CRC64;

  Query Match             15.8%;  Score 262;  DB 1;  Length 382;
  Best Local Similarity   25.1%;  Pred. No. 9.3e-11;
  Matches   88;  Conservative   76;  Mismatches  122;  Indels   64;  Gaps   14;

Qy          1 MDPTISTLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNA 60
              : |    |:  :   | :  :| |   :|: :|  : ||:||  | ::|:::|||||:
Db          6 LKPLCPLLEEMVQLPNHSNSSLRYIDHVSV-LLHGLASLLGLVENGLILFVVGCRMRQTV 64

Qy         61 FSIYILNLAAADFLFLSGRLIYSLLSFISIPH------TISKILYPVMMFSYFAGLSFLS 114
               : ::|:|| :| |  :    ::   |::: |      | |:  |   : ||    ||
Db         65 VTTWVLHLALSDLLAAASLPFFTY--FLAVGHSWELGTTFCKLHSSVFFLNMFASGFLLS 122

Qy        115 AVSTERCLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEW-----------MLCG 162
              |:| :||| |: |:| : ||   ::  ||::||||::| :|  :           ::|
Db        123 AISLDRCLQVVRPVWAQNHRTVAVAHRVCLMLWALAVLNTIPYFVFRDTIPRLDGRIMCY 182

Qy        163 FLF----SGAD-SAWCQ-------TSDFITVAWLIFLCVVLCGSSLVLLIRILCGSRKIP 210
              :       | |    |          | |: :|::: | ::   | : : :|:    |:
Db        183 YNLLLWNPGPDRDTTCDYRQKALAVSKFL-LAFMVPLAII-ASSHVAVSLRLHHRGRQ-- 238

Qy        211 LTRLYVTILLTVLV-FLLCGLPFGIQFFLFLWIHVDREVLFCHVHLVSIF---------- 259
                |  :|  ::  ::| |:||  |: |                :|    | |:
Db        239 RTGRFVRLVAAIVVAFVLCWGPYHI-----------FSLLEARAHSVTTLRQLASRGLPF 287

Qy        260 ---LSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASEVDEGGG 306
                 |:  ||  ||::| |       : |::|: ||:  | : |  |: ||
Db        288 VTSLAFFNSVVNPLLYVFTCPDMLYKLRRSLRAVLESVLVEDS--DQSGG 335


RESULT 8
CML1_HUMAN
ID   CML1_HUMAN     STANDARD;      PRT;   373 AA.
AC   Q99788; Q99789; O75748;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Chemokine receptor-like 1 (G-protein coupled receptor DEZ) (G protein-
DE   coupled receptor ChemR23).
GN   CMKLR1 or DEZ or CHEMR23.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX   MEDLINE=97289630; PubMed=9144535;
RA   Methner A., Hermey G., Schinke B., Hermans-Borgmeyer I.;
RT   "A novel G protein-coupled receptor with homology to neuropeptide and
RT   chemoattractant receptors expressed during bone development.";
RL   Biochem. Biophys. Res. Commun. 233:336-342(1997).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM B).
RX   MEDLINE=98264639; PubMed=9603476;
RA   Samson M., Edinger A.L., Stordeur P., Rucker J., Verhasselt V.,
RA   Sharron M., Govaerts C., Mollereau C., Vassart G., Doms R.W.,
RA   Parmentier M.;
RT   "ChemR23, a putative chemoattractant receptor, is expressed in
RT   monocyte-derived dendritic cells and macrophages and is a coreceptor
RT   for SIV and some primary HIV-1 strains.";
RL   Eur. J. Immunol. 28:1689-1700(1998).
CC   -!- FUNCTION: ORPHAN RECEPTOR. COULD BE A CHEMOTACTIC PEPTIDE
CC       RECEPTOR. MAY HAVE A FUNCTION IN BONE METABOLISM. ACTS AS A
CC       CORECEPTOR FOR SEVERAL SIV STRAINS (SIVMAC316, SIVMAC239,
CC       SIVMACL7E-FR AND SIVSM62A), AS WELL AS A PRIMARY HIV-1 STRAIN
CC       (92UG024-2).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A (SHOWN HERE) AND B; ARE
CC       PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: PROMINENTLY EXPRESSED IN DEVELOPING OSSEOUS
CC       AND CARTILAGINOUS TISSUE. ALSO FOUND IN ADULT PARATHYROID GLANDS.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED DURING BONE DEVELOPMENT.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U79526; AAC51258.1; -.
DR   EMBL; U79527; AAC51259.1; -.
DR   EMBL; Y14838; CAA75112.1; -.
DR   Genew; HGNC:2121; CMKLR1.
DR   MIM; 602351; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 2.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Alternative splicing.
FT   DOMAIN       1     41       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    42     64       1 (POTENTIAL).
FT   DOMAIN      65     75       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    76     97       2 (POTENTIAL).
FT   DOMAIN      98    114       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   115    135       3 (POTENTIAL).
FT   DOMAIN     136    154       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   155    176       4 (POTENTIAL).
FT   DOMAIN     177    224       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   225    245       5 (POTENTIAL).
FT   DOMAIN     246    261       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   262    282       6 (POTENTIAL).
FT   DOMAIN     283    300       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   301    320       7 (POTENTIAL).
FT   DOMAIN     321    373       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD     9      9       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   192    192       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DISULFID   112    189       POTENTIAL.
FT   VARSPLIC     1      2       MISSING (IN ISOFORM B).
FT   CONFLICT   248    248       Q -> H (IN REF. 1; AAC51258).
SQ   SEQUENCE   373 AA;  42322 MW;  5244B9738EC93834 CRC64;

  Query Match             15.4%;  Score 256.5;  DB 1;  Length 373;
  Best Local Similarity   25.8%;  Pred. No. 2.1e-10;
  Matches   90;  Conservative   63;  Mismatches  121;  Indels   75;  Gaps   14;

Qy          5 ISTLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNAFSIY 64
              :| |:  :| |                | |:  ||  :|: || :|: :   :|::    ::
Db         31 LSPLEARVTRI------------FLVVVYSIVCFLGILGNGLVIIIATFKMKKTVNMVW 77

Qy         65 ILNLAAADFL---FLSGRLIYSLLSFISIPHT-ISKILYPVMMFSYFAGLSFLSAVSTER 120
               |||| ||||   ||   : |: :  :  | : ||   ::: : |  :  |: :|::|
Db         78 FLNLAVADFLFNVFLPIHITYAAMDYHWVFGTAMCKISNFLLIHNMFTSVFLLTIISSDR 137

Qy        121 CLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSI----------LEWMLCGF----LFS 166
              |:||| |:| : ||   |: : |:::| |:   |              |   :  |    | :
Db        138 CISVLLPVWSQNHRSVRLAYMACMVIWVLAFFLSSPSLVFRDTANLHGKISCFNNFSLST 197

Qy        167 GADSAWCQTSDFITVAWLIFLCVV----LCGSSLVLLI------RILCGSRKIPLTRLYV 216
                |:|  |    | :   : |     |||  : :||       |:|   |:   ||
Db        198 PGSSSWPTHSQMDPVGYSRHMVVTVTRFLCGFLVPVLIITACYLTIVC---KLQRNRLAK 254

Qy        217 T-----ILLTVLV-FLLCGLPFGIQFFLFLWIHVDREVLFCHVHLV-SIF---------L 260
              |     |::|::: | ||  |:          |   :   |  : :|:|          |
Db        255 TKKPFKIIVTIIITFFLCWCPY----------HTLNLLELHHTAMPGSVFSLGLPLATAL 304

Qy        261 SALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASEVDEGGGQLP 309
              :  ||  |||:| |:|     |: :  |: |   | :|   | |    |
Db        305 AIANSCMNPILYVFMG-----QDFKKFKVALFSRLVNALSEDTGHSSYP 348
```

```
RESULT 9
FML2_PANTR
ID   FML2_PANTR .   STANDARD;      PRT;   349 AA.
AC   P79243;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   N-formyl peptide receptor-like 2 receptor (Fragment).
GN   FPRL2.
OS   Pan troglodytes (Chimpanzee).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX   NCBI_TaxID=9598;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96421539; PubMed=8824156;
RA   Alvarez V., Coto E., Sehen F., Gouzalek-Koces S., Lopez-Larrea C.;
RT   "Molecular evolution of the N-formyl peptide and C5a receptors in
RT   non-human primates.";
RL   Immunogenetics 44:446-452(1996).
CC   -!- FUNCTION: LOW AFFINITY RECEPTOR TO N-FORMYL-METHIONYL PEPTIDES,
CC       WHICH ARE POWERFUL NEUTROPHILS CHEMOTACTIC FACTORS. BINDING OF
CC       FMLP TO THE RECEPTOR CAUSES ACTIVATION OF NEUTROPHILS. THIS
CC       RESPONSE IS MEDIATED VIA A G-PROTEIN THAT ACTIVATES A
CC       PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X97743; CAA66327.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Chemotaxis.
FT   DOMAIN        1     27       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     28     50       1 (POTENTIAL).
FT   DOMAIN       51     61       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     62     83       2 (POTENTIAL).
FT   DOMAIN       84    100       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    101    121       3 (POTENTIAL).
FT   DOMAIN      122    140       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    141    162       4 (POTENTIAL).
FT   DOMAIN      163    205       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    206    226       5 (POTENTIAL).
FT   DOMAIN      227    242       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    243    266       6 (POTENTIAL).
FT   DOMAIN      267    286       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    287    306       7 (POTENTIAL).
FT   DOMAIN      307   >349       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD      4      4       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     10     10       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DISULFID     98    176       POTENTIAL.
FT   NON_TER     349    349
SQ   SEQUENCE   349 AA;  39611 MW;  88279F8C72915383 CRC64;

  Query Match             15.4%;  Score 256;  DB 1;  Length 349;
  Best Local Similarity   27.6%;  Pred. No. 2.1e-10;
  Matches   92;  Conservative   60;  Mismatches  125;  Indels   56;  Gaps   13;

Qy         14 PINGTEETLCYKQTLSLTVLTCIVSLV-------GLTGNAVVLWLLGCRMRRNAFSIYIL 66
              |:| ||| |   :    |||     ||       |: || :|:|: | || |   :|  |
Db          8 PLNETEEVL--PEPAGHTVLWIFSLLVHGVTFVFGVLGNGLVIWVAGFRMTRTVNTICYL 65

Qy         67 NLAAADFLFLSGRLIYSLLS--------FISIPHTISKILYPVMMFSYFAGLSFLSAVST 118
              ||| ||| | |  | : ::|         | |    : |::: ::  : |  :  :: ::
Db         66 NLALADFSF-SAILPFRMVSVAMREKWPFGSF---LCKLVHVMIDINLFVSVYLITIIAL 121

Qy        119 ERCLSVLWPIWYRCHRPTHLSAVVCVLLWALSL---LRSILEW-----------MLCGFL 164
             :||: || | | : ||    |:  |    || |::   | : : |           :  |
Db        122 DRCICVLHPAWAQNHRTMSLAKRVMTGLWILTIVLTLPNFIFWTTIRTTNGDTYCIFNFA 181

Qy        165 FSGADSAWCQTSDFITVAWLIFLCVVLCGSSLVLLIRILC---------GSRKIPLTRLY 215
             | | |:|  : : |||:| :  :   :  | |: : |  :|          :  |  :|
Db        182 FWG-DTAVERLNVFITMAKVFLILHFIIGFSMPMSIITVCYGIIAAKIHRNHMIKSSRPL 240

Qy        216 VTILLTVLVFLLCGLPFG-IQFFLFLWIHVDREVLF---CHVHLVSI----FLSALNSSA 267
                 |  | :|  |:  |   : :|:   :|:|      : || |     |:  ||
Db        241 RVFAAVVASFFICWFPYELIGILMAVWL---KEMLLNGKYKIILVLINPTSSLAFFNSCL 297

Qy        268 NPIIYFFVGSFRQRQNRQNLKLVLQRALQDASE 300
             |||:| |:|   | :  ::|   |:||| :  :
Db        298 NPILYVFMGRNFQERLIRSLPTSLERALTEVPD 330


RESULT 10
C5AR_GORGO
ID   C5AR_GORGO     STANDARD;      PRT;   340 AA.
AC   P79175;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   C5a anaphylatoxin chemotactic receptor (C5a-R) (Fragment).
GN   C5R1 OR C5AR.
OS   Gorilla gorilla gorilla (Lowland gorilla).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
OX   NCBI_TaxID=9595;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96421539; PubMed=8824156;
RA   Alvarez V., Coto E., Sehen F., Gouzalek-Koces S., Lopez-Larrea C.;
RT   "Molecular evolution of the N-formyl peptide and C5a receptors in
RT   non-human primates.";
RL   Immunogenetics 44:446-452(1996).
CC   -!- FUNCTION: RECEPTOR FOR THE CHEMOTACTIC AND INFLAMMATORY PEPTIDE
CC       ANAPHYLATOXIN C5A. THIS RECEPTOR STIMULATES CHEMOTAXIS, GRANULE
CC       ENZYME RELEASE AND SUPEROXIDE ANION PRODUCTION.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- PTM: SULFATION PLAYS A CRITICAL ROLE IN THE ASSOCIATION OF C5AR
CC       WITH C5A, BUT NO SIGNIFICANT ROLE IN THE ABILITY OF THE RECEPTOR
CC       TO TRANSDUCE A SIGNAL AND MOBILIZE CALCIUM IN RESPONSE TO A SMALL
CC       PEPTIDE AGONIST (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X97733; CAA66317.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation;
KW   Chemotaxis.
FT   NON_TER       1      1
FT   DOMAIN       <1     30       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     31     53       1 (POTENTIAL).
FT   DOMAIN       54     64       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     65     87       2 (POTENTIAL).
FT   DOMAIN       88    103       EXTRACELLULAR (POTENTIAL).
```

```
FT   TRANSMEM     104     125       3 (POTENTIAL).
FT   DOMAIN       126     146       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     147     167       4 (POTENTIAL).
FT   DOMAIN       168     193       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     194     219       5 (POTENTIAL).
FT   DOMAIN       220     235       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     236     258       6 (POTENTIAL).
FT   DOMAIN       259     275       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     276     296       7 (POTENTIAL).
FT   DOMAIN       297    >340       CYTOPLASMIC (POTENTIAL).
FT   DISULFID     102     181       BY SIMILARITY.
FT   MOD_RES        4       4       SULFATION (BY SIMILARITY).
FT   MOD_RES        7       7       SULFATION (BY SIMILARITY).
FT   NON_TER      340     340
SQ   SEQUENCE   340 AA;  38235 MW;  214CFCE78A47B304 CRC64;

  Query Match             15.2%;  Score 252;  DB 1;  Length 340;
  Best Local Similarity   26.6%;  Pred. No. 3.8e-10;
  Matches   89;  Conservative   57;  Mismatches  129;  Indels   60;  Gaps   10;

Qy          7 TLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNAFSIYIL 66
              |||   ||:: |  ||      |:| |:  :| ||| ||||:|:|:     :|   :|: |
Db         12 TLDPN-TPVDKTSNTLRVPDILAL-VIFAVVFLVGVLGNAMVVWVTAFEAKRTINAIWFL 69

Qy         67 NLAAADFL-FLSGRLIYSLLSFISIPH-----TISKILYPVMMFSYFAGLSFLSAVSTER 120
              ||| ||||  |:  :::: . | :   |          :||  ::: : :| :  |: :| :|
Db         70 NLAVADFLSCLALPILFT--SIVQHHHWPFGGAACRILPSLILLNMYASILLLATISADR 127

Qy        121 CLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQTSDFIT 180
              | |  ||| :  |   |: : | : | |:|| :|  ::                 :
Db        128 FLLVFKPIWCQNFRGAGLAWIACAVAWGLALLLTIPSFL-----------------YRV 169

Qy        181 VAWLIFLCVVLCG---------SSLVLLIRILCGS--RKIPLTRLYVTILL--------- 220
              |    |  ||||           | ::|:: |     : ||  |  |||
Db        170 VREEYFPPKVLCGVDYSHDKRRERAVAIVRLVLGFLWPLLTLTICYTFILLRTWSRRATR 229

Qy        221 ----------TVLVFLLCGLPFGIQFFLFLWIHVDREVLFCHVHLVSIFLS--ALNSSAN 268
                       |  | :  ||: :   :  ::             | |: :|   :|   |
Db        230 STKTLKVVVAVVASFFIFWLPYQVTGIMMSFLEPSSPTFLLLNKLDSLCVSFAYINCCIN 289

Qy        269 PIIYFFVGSFRQRQNRQNLKLVLQRALQDASEVDE 303
              ||||    |   | : |::|  :|:  | : | | |
Db        290 PIIYVVAGQGFQGRLRKSLPSLLRNVLTEESVVRE 324


RESULT 11
FML2_GORGO
ID   FML2_GORGO     STANDARD;      PRT;   349 AA.
AC   P79178;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   FMLP-related receptor II (FMLP-R-II) (Fragment).
GN   FPRL2.
OS   Gorilla gorilla gorilla (Lowland gorilla).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
OX   NCBI_TaxID=9595;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96421539; PubMed=8824156;
RA   Alvarez V., Coto E., Sehen F., Gouzalek-Koces S., Lopez-Larrea C.;
RT   "Molecular evolution of the N-formyl peptide and C5a receptors in
RT   non-human primates.";
RL   Immunogenetics 44:446-452(1996).
CC   -!- FUNCTION: LOW AFFINITY RECEPTOR TO N-FORMYL-METHIONYL PEPTIDES,
CC       WHICH ARE POWERFUL NEUTROPHILS CHEMOTACTIC FACTORS. BINDING OF
CC       FMLP TO THE RECEPTOR CAUSES ACTIVATION OF NEUTROPHILS. THIS
CC       RESPONSE IS MEDIATED VIA A G-PROTEIN THAT ACTIVATES A
CC       PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ----------------------------------------------------------------------
```

```
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X97742; CAA66326.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Chemotaxis.
FT   DOMAIN       1      27       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    28      50       1 (POTENTIAL).
FT   DOMAIN      51      61       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    62      83       2 (POTENTIAL).
FT   DOMAIN      84     100       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   101     121       3 (POTENTIAL).
FT   DOMAIN     122     140       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   141     162       4 (POTENTIAL).
FT   DOMAIN     163     205       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   206     226       5 (POTENTIAL).
FT   DOMAIN     227     242       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   243     266       6 (POTENTIAL).
FT   DOMAIN     267     286       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   287     306       7 (POTENTIAL).
FT   DOMAIN     307    >349       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD     4       4       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    10      10       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DISULFID    98     176       POTENTIAL.
FT   NON_TER    349     349
SQ   SEQUENCE   349 AA;  39432 MW;  C6ED77CFDE023834 CRC64;

  Query Match             15.2%;  Score 252;  DB 1;  Length 349;
  Best Local Similarity   27.6%;  Pred. No. 3.9e-10;
  Matches   92;  Conservative   60;  Mismatches  125;  Indels   56;  Gaps   13;

Qy         14 PINGTEETLCYKQTLSLTVLTCIVSLV-------GLTGNAVVLWLLGCRMRRNAFSIYIL 66
              |:| ||| |   :    |||     ||        |: || :|:|: |  | |  :|  |
Db          8 PLNETEEVL--PEPAGHTVLWIFSLLVHGVTFIFGVLGNGLVIWVAGFLMTRTVNTICYL 65

Qy         67 NLAAADFLFLSGRLIYSLLS--------FISIPHTISKILYPVMMFSYFAGLSFLSAVST 118
              ||| ||| | |  | : ::|        | |    : |::: ::  : |  :  :: ::
Db         66 NLALADFSF-SAILPFHMVSVAMREKWPFGSF---LCKLVHVMIDINLFVSVYLITIIAL 121

Qy        119 ERCLSVLWPIWYRCHRPTHLSAVVCVLLWALSL---LRSILEW----------MLCGFL 164
              :||: || | | : ||    |:  |    || |::   | : : |           :  |
Db        122 DRCICVLHPAWAQNHRTMSLAKRVMTGLWILTIVLTLPNFIFWTTISTTNGDTYCIFNFP 181

Qy        165 FSGADSAWCQTSDFITVAWLIFLCVVLCGSSLVLLIRILC---------GSRKIPLTRLY 215
              | | |:|  : : |||:| :  :   : | |: : |  :|          :  | :|
Db        182 FWG-DTAVERLNVFITMAKVFLILHFIIGFSMPMSIITVCYGIIAAKIHRNHMIKSSRPL 240

Qy        216 VTILLTVLVFLLCGLPFG-IQFFLFLWIHVDREVLF---CHVHLVSI----FLSALNSSA 267
                  |  | :|  |:  |   : :|:   :|:|      : || |     |:  ||
Db        241 RVFAAVVASFFICWFPYELIGILMAVWL---KEMLLNGKYKIILVLINPTSSLAFFNSCL 297

Qy        268 NPIIYFFVGSFRQRQNRQNLKLVLQRALQDASE 300
              |||:| |:||  | :  ::|   |:||| :  :
Db        298 NPILYVFLGSNFQERLIRSLPTSLERALTEVPD 330


RESULT 12
FML2_PONPY
ID   FML2_PONPY     STANDARD;      PRT;   349 AA.
AC   P79237;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   N-formyl peptide receptor-like 2 receptor (Fragment).
```

```
GN   FPRL2.
OS   Pongo pygmaeus (Orangutan).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
OX   NCBI_TaxID=9600;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96421539; PubMed=8824156;
RA   Alvarez V., Coto E., Sehen F., Gouzalek-Koces S., Lopez-Larrea C.;
RT   "Molecular evolution of the N-formyl peptide and C5a receptors in
RT   non-human primates.";
RL   Immunogenetics 44:446-452(1996).
CC   -!- FUNCTION: LOW AFFINITY RECEPTOR TO N-FORMYL-METHIONYL PEPTIDES,
CC       WHICH ARE POWERFUL NEUTROPHILS CHEMOTACTIC FACTORS. BINDING OF
CC       FMLP TO THE RECEPTOR CAUSES ACTIVATION OF NEUTROPHILS. THIS
CC       RESPONSE IS MEDIATED VIA A G-PROTEIN THAT ACTIVATES A
CC       PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X97741; CAA66325.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Chemotaxis.
FT   DOMAIN        1     27       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     28     50       1 (POTENTIAL).
FT   DOMAIN       51     61       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     62     83       2 (POTENTIAL).
FT   DOMAIN       84    100       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    101    121       3 (POTENTIAL).
FT   DOMAIN      122    140       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    141    162       4 (POTENTIAL).
FT   DOMAIN      163    205       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    206    226       5 (POTENTIAL).
FT   DOMAIN      227    242       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    243    266       6 (POTENTIAL).
FT   DOMAIN      267    286       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    287    306       7 (POTENTIAL).
FT   DOMAIN      307   >349       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD      4      4       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     10     10       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DISULFID     98    176       POTENTIAL.
FT   NON_TER     349    349
SQ   SEQUENCE   349 AA;  39423 MW;  C8298D223395EECB CRC64;

  Query Match             15.2%;  Score 252;  DB 1;  Length 349;
  Best Local Similarity   26.5%;  Pred. No. 3.9e-10;
  Matches   93;  Conservative   65;  Mismatches  129;  Indels   64;  Gaps   15;

Qy         14 PINGTEETLCYKQTLSLTVLTCIVSLV-------GLTGNAVVLWLLGCRMRRNAFSIYIL 66
              |:| :|| |   :    |||     ||       |: || :|:|: | || |   :|  |
Db          8 PLNESEEVL--PEPAGHTVLWIFSLLVHGVTFIFGVLGNGLVIWVAGFRMTRTVNTICYL 65

Qy         67 NLAAADFLFLSGRLIYSLLSFI-----SIPHTISKILYPVMMFSYFAGLSFLSAVSTERC 121
              ||| ||| | |  | : ::|           : |::: ::  : |  :|  :: ::|:||
Db         66 NLALADFSF-SAILPFRMVSVAMREKWPFGTFLCKLVHVMIDINLFVSVYLITIIALDRC 124

Qy        122 LSVLWPIWYRCHRPTHLSAVVCVLLWALSL---LRSILEW----------MLCGFLFSG 167
              : || | | : ||   |:  | : || |::   | : : |          :  | | |
Db        125 ICVLHPAWAQNHRTMSLAKRVMMGLWILAIVLTLPNFIFWTTISTKNGDTYCIFNFPFWG 184

Qy        168 ADSAWCQTSDFITVAWLIFLCVVLCGSSLVLLIRILCG---SRKI----------PLTRL 214
```

```
              |:|  : : |||:  :  :   : | |: : |  :|    : ||          || |:
Db        185 -DTAVERLNAFITMGKVFLILHFIIGFSMPMSIITVCYGIIAAKIHRNHMIKSSSPL-RV 242

Qy        215 YVTILLTVLVFLLCGLPFG-IQFFLFLWIHVDREVLFCHVHLVSIF-------LSALNSS 266
              :  :   |  | :|  |:  |   : :|:  :|:|    : : :         |:  ||
Db        243 FAAV---VASFFICWFPYELIGILMAVWL---KEMLLNGKYKIILVLLNPTSSLAFFNSC 296

Qy        267 ANPIIYFFVGSFRQRQNRQNLKLVLQRALQD------ASEVDEGGGQLPEE 311
              |||:| |:||  | :  ::|   |:||| :       |          |||
Db        297 LNPILYVFLGSNFQERLIRSLPTSLERALTEVPDSAQTSNTHTNSASPPEE 347


RESULT 13
C3AR_RAT
ID   C3AR_RAT       STANDARD;      PRT;   473 AA.
AC   O55197;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   C3a anaphylatoxin chemotactic receptor (C3a-R) (C3AR).
GN   C3AR1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Wistar; TISSUE=Lung;
RX   MEDLINE=98125550; PubMed=9464274;
RA   Fukuoka Y., Ember J.A., Hugli T.E.;
RT   "Cloning and characterization of rat C3a receptor: differential
RT   expression of rat C3a and C5a receptors by LPS stimulation.";
RL   Biochem. Biophys. Res. Commun. 242:663-668(1998).
CC   -!- FUNCTION: RECEPTOR FOR THE CHEMOTACTIC AND INFLAMMATORY PEPTIDE
CC       ANAPHYLATOXIN C3A. THIS RECEPTOR STIMULATES CHEMOTAXIS, GRANULE
CC       ENZYME RELEASE AND SUPEROXIDE ANION PRODUCTION.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U86379; AAC40071.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Chemotaxis.
FT   DOMAIN        1     23       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     24     46       1 (POTENTIAL).
FT   DOMAIN       47     57       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     58     80       2 (POTENTIAL).
FT   DOMAIN       81     96       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     97    118       3 (POTENTIAL).
FT   DOMAIN      119    139       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    140    160       4 (POTENTIAL).
FT   DOMAIN      161    329       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    330    349       5 (POTENTIAL).
FT   DOMAIN      350    366       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    367    389       6 (POTENTIAL).
FT   DOMAIN      390    406       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    407    427       7 (POTENTIAL).
FT   DOMAIN      428    473       CYTOPLASMIC (POTENTIAL).
FT   DISULFID     95    172       BY SIMILARITY.
FT   CARBOHYD      9      9       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    201    201       N-LINKED (GLCNAC. . .) (POTENTIAL).
```

```
SQ   SEQUENCE   473 AA;  52896 MW;  722F16F6F4B16451 CRC64;

  Query Match             15.2%;  Score 252;  DB 1;  Length 473;
  Best Local Similarity   20.5%;  Pred. No. 5.1e-10;
  Matches   95;  Conservative   66;  Mismatches  112;  Indels  190;  Gaps   14;

Qy          7 TLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNAFSIYIL 66
              | ||  | ::     |   | ::  |:  :  |:|| || :|||: | :|:|   ::: |
Db          5 TADTNSTDLH--SRPLFKPQDIASMVILSLTCLLGLPGNGLVLWVAGVKMKRTVNTVWFL 62

Qy         67 NLAAADF-----------LFLSGRLIYSLLSFISIPHTISKILYPVMMFSYFAGLSFLS 114
              :|  |||           | | |   | |         : |::  |:: : || :  |:
Db         63 HLTLADFLCCLSLPFSVAHLILRGHWPYGLF--------LCKLIPSVIILNMFASVFLLT 114

Qy        115 AVSTERCLSVLWPIWYRCHRPTHLSAVVCVLLWALSLL--------RSIL---EWMLCGF 163
              |:| :||| |  ||| : ||    :  ||  :| :: :        | :|   :: :||:
Db        115 AISLDRCLMVHKPIWCQNHRSVRTAFAVCGCVWVVTFVMCIPVFVYRDLLVVDDYSVCGY 174

Qy        164 LFSG----------------------------------------ADSAWCQTS-------- 176
              |                                          |  |  |:
Db        175 NFDSSRAYDYWDYMYNSHLPEINPPDNSTGHVDDRTAPSSSVPARDLWTATTALQSQTFH 234

Qy        177 ------------------------------------------------------------ 176

Db        235 TSPEDPFSQDSASQQPHYGGKPPTVLIATIPGGFPVEDHKSNTLNTGAFLSAHTEPSLTA 294

Qy        177 --------DF--------------------ITVAWLI------FLCVVLCGSSLVLLIRI 202
                     ||                    ||:: |:      |  :: | | :|  :
Db        295 SSSPLYAHDFPDDYFDQLMYGNHAWTPQVAITISRLVVGFLVPFFIMITCYSLIVFRM-- 352

Qy        203 LCGSRKIPLTR-----LYVTILLTVLVFLLCGLPFGIQFFLFLWIHVD---REVLFCHVH 254
                ||  ||:     | | : : | || :| :|: |   | :    :   |||:    |
Db        353 ----RKTNLTKSRNKTLRVAVAV-VTVFFVCWIPYHIVGILLVITDQESALREVVLPWDH 407

Qy        255 LVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQD 297
              : || |:: ||  || :|  :|   ::: ||::| :|: |  :
Db        408 M-SIALASANSCFNPFLYALLGKDFRKKARQSVKGILEAAFSE 449


RESULT 14
FML2_HUMAN
ID   FML2_HUMAN     STANDARD;      PRT;   353 AA.
AC   P25089;
DT   01-MAY-1992 (Rel. 22, Created)
DT   01-MAY-1992 (Rel. 22, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   FMLP-related receptor II (FMLP-R-II).
GN   FPRL2 OR FPRH1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92307681; PubMed=1612600;
RA   Bao L., Gerard N.P., Eddy R.L. Jr., Shows T.B., Gerard C.;
RT   "Mapping of genes for the human C5a receptor (C5AR), human FMLP
RT   receptor (FPR), and two FMLP receptor homologue orphan receptors
RT   (FPRH1, FPRH2) to chromosome 19.";
RL   Genomics 13:437-440(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94256976; PubMed=8198572;
RA   Durstin M., Gao J.-L., Tiffany H.L., McDermott D., Murphy P.M.;
RT   "Differential expression of members of the N-formylpeptide receptor
RT   gene cluster in human phagocytes.";
RL   Biochem. Biophys. Res. Commun. 201:174-179(1994).
CC   -!- FUNCTION: LOW AFFINITY RECEPTOR TO N-FORMYL-METHIONYL PEPTIDES,
CC       WHICH ARE POWERFUL NEUTROPHILS CHEMOTACTIC FACTORS. BINDING OF
CC       FMLP TO THE RECEPTOR CAUSES ACTIVATION OF NEUTROPHILS. THIS
CC       RESPONSE IS MEDIATED VIA A G-PROTEIN THAT ACTIVATES A
CC       PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M76673; AAA58482.1; -.
DR   EMBL; L14061; AAA52474.1; -.
DR   PIR; C42009; C42009.
DR   Genew; HGNC:3828; FPRL2.
DR   MIM; 136539; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Chemotaxis.
FT   DOMAIN        1     27       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     28     50       1 (POTENTIAL).
FT   DOMAIN       51     61       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     62     83       2 (POTENTIAL).
FT   DOMAIN       84    100       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    101    121       3 (POTENTIAL).
FT   DOMAIN      122    140       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    141    162       4 (POTENTIAL).
FT   DOMAIN      163    205       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    206    226       5 (POTENTIAL).
FT   DOMAIN      227    242       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    243    266       6 (POTENTIAL).
FT   DOMAIN      267    286       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    287    306       7 (POTENTIAL).
FT   DOMAIN      307    353       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD      4      4       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     10     10       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DISULFID     98    176       POTENTIAL.
FT   CONFLICT     94     94       A -> G (IN REF. 2).
FT   CONFLICT    211    211       T -> S (IN REF. 2).
FT   CONFLICT    338    338       H -> D (IN REF. 2).
SQ   SEQUENCE   353 AA;  40015 MW;  8ED7450A14A36C9A CRC64;

  Query Match             15.1%;  Score 251;  DB 1;  Length 353;
  Best Local Similarity   27.0%;  Pred. No. 4.6e-10;
  Matches   90;  Conservative   61;  Mismatches  126;  Indels   56;  Gaps   13;

Qy         14 PINGTEETLCYKQTLSLTVLTCIVSLV-------GLTGNAVVLWLLGCRMRRNAFSIYIL 66
              |:| ||| |   :     |||      ||       |: || :|:|: | || |   :|  |
Db          8 PLNETEEVL--PEPAGHTVLWIFSLLVHGVTFVFGVLGNGLVIWVAGFRMTRTVNTICYL 65

Qy         67 NLAAADFLFLSGRLIYSLLS--------FISIPHTISKILYPVMMFSYFAGLSFLSAVST 118
              ||| ||| | |   | : ::|         | |    : |::: ::  : |  :  :: ::
Db         66 NLALADFSF-SAILPFRMVSVAMREKWPFASF---LCKLVHVMIDINLFVSVYLITIIAL 121

Qy        119 ERCLSVLWPIWYRCHRPTHLSAVVCVLLWALSL---LRSILEW----------MLCGFL 164
              :||: || | | : ||   |:  |   ||  ::   | : : |              : |
Db        122 DRCICVLHPAWAQNHRTMSLAKRVMTGLWIFTIVLTLPNFIFWTTISTTNGDTYCIFNFA 181

Qy        165 FSGADSAWCQTSDFITVAWLIFLCVVLCGSSLVLLIRILC---------GSRKIPLTRLY 215
              | | |:|  : : |||:| :  :   : | :: : |  :|           :  |  :|
Db        182 FWG-DTAVERLNVFITMAKVFLILHFIIGFTVPMSIITVCYGIIAAKIHRNHMIKSSRPL 240

Qy        216 VTILLTVLVFLLCGLPFG-IQFFLFLWIHVDREVLF---CHVHLVSI----FLSALNSSA 267
                    | | :|  |:  |   : :|:   :|:|       : || |     |:  ||
Db        241 RVFAAVVASFFICWFPYELIGILMAVWL---KEMLLNGKYKIILVLINPTSSLAFFNSCL 297

Qy        268 NPIIYFFVGSFRQRQNRQNLKLVLQRALQDASE 300
              |||:| |:|   | :  ::|   |:||| :  :
Db        298 NPILYVFMGRNFQERLIRSLPTSLERALTEVPD 330
```

```
RESULT 15
C5AR_MACMU
ID   C5AR_MACMU     STANDARD;      PRT;   340 AA.
AC   P79188;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   C5a anaphylatoxin chemotactic receptor (C5a-R) (Fragment).
GN   C5R1 OR C5AR.
OS   Macaca mulatta (Rhesus macaque).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9544;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96421539; PubMed=8824156;
RA   Alvarez V., Coto E., Sehen F., Gouzalek-Koces S., Lopez-Larrea C.;
RT   "Molecular evolution of the N-formyl peptide and C5a receptors in
RT   non-human primates.";
RL   Immunogenetics 44:446-452(1996).
CC   -!- FUNCTION: RECEPTOR FOR THE CHEMOTACTIC AND INFLAMMATORY PEPTIDE
CC       ANAPHYLATOXIN C5A. THIS RECEPTOR STIMULATES CHEMOTAXIS, GRANULE
CC       ENZYME RELEASE AND SUPEROXIDE ANION PRODUCTION.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- PTM: SULFATION PLAYS A CRITICAL ROLE IN THE ASSOCIATION OF C5AR
CC       WITH C5A, BUT NO SIGNIFICANT ROLE IN THE ABILITY OF THE RECEPTOR
CC       TO TRANSDUCE A SIGNAL AND MOBILIZE CALCIUM IN RESPONSE TO A SMALL
CC       PEPTIDE AGONIST (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X97731; CAA66315.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation;
KW   Chemotaxis.
FT   NON_TER       1      1
FT   DOMAIN       <1     30       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     31     53       1 (POTENTIAL).
FT   DOMAIN       54     64       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     65     87       2 (POTENTIAL).
FT   DOMAIN       88    103       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    104    125       3 (POTENTIAL).
FT   DOMAIN      126    146       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    147    167       4 (POTENTIAL).
FT   DOMAIN      168    193       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    194    219       5 (POTENTIAL).
FT   DOMAIN      220    235       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    236    258       6 (POTENTIAL).
FT   DOMAIN      259    275       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    276    296       7 (POTENTIAL).
FT   DOMAIN      297   >340       CYTOPLASMIC (POTENTIAL).
FT   DISULFID    102    181       BY SIMILARITY.
FT   MOD_RES       4      4       SULFATION (BY SIMILARITY).
FT   MOD_RES       7      7       SULFATION (BY SIMILARITY).
FT   NON_TER     340    340
SQ   SEQUENCE   340 AA;  38274 MW;  E11F7C73AB97FFBB CRC64;

  Query Match           15.1%;  Score 250;  DB 1;  Length 340;
  Best Local Similarity  26.2%;  Pred. No. 5.2e-10;
  Matches   85;  Conservative   66;  Mismatches  135;  Indels   38;  Gaps   11;

Qy          7 TLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNAFSIYIL 66
              |||   ||:: |  ||     |:| |:  :| |||:  ||:|:|:     :|   :|: |
Db         12 TLDAN-TPVDKTSNTLRVPDILAL-VIFAVVFLVGVLRNALVVWVTAFEAKRTINAIWFL 69

Qy         67 NLAAADFL-FLSGRLIYSLLSFISIPH-----TISKILYPVMMFSYFAGLSFLSAVSTER 120
              ||| |||| |:  :::: | :   |        :||  ::: : :| :  |: :| :|
Db         70 NLAVADFLSCLALPILFT--SIVQHHHWPFGGAACRILPSLILLNMYASILLLATISADR 127

Qy        121 CLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWM------------LCGFLFSG 167
              | |  ||| :  |   |: : | : | |:|| :|  ::             ||| :  |
Db        128 FLLVFNPIWCQNFRGAGLAWIACAVAWGLALLLTIPSFLYRVVREEYFPPKVLCG-VDHG 186

Qy        168 ADSAWCQTSDFITVAWLI------FLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLT 221
              |    :   : :| |:       | : :| : |:|     :|    :: | :: :
Db        187 HDK---RRERAVAIARLVLGFVWPLLTLTMCYTFLLLRTWSRRATRSTKTLKVVVAVVAS 243

Qy        222 VLVFLLCGLPFGIQFFLFLWIHVDREVLFCHVHLVSIFLS--ALNSSANPIIYFFVGSFR 279
              :|    ||: :   :  ::           | |: :|   :|   |||||   |
Db        244 FFIF---WLPYQVTGMMMSFLEPSSPTFLLLKKLDSLCISFAYINCCINPIIYVVAGQGF 300

Qy        280 QRQNRQNLKLVLQRALQDASEVDE 303
              | : |::|  :|:  | : | | |
Db        301 QGRLRKSLPSLLRNVLTEESMVRE 324


Search completed: July  2, 2003, 19:21:01
Job time : 12 secs
```

GenCore version 5.1.6


OM protein - protein search, using sw model

Run on: July 2, 2003, 19:18:35 ; Search time 19 Seconds
(without alignments)
1629.225 Million cell updates/sec

Title: US-09-849-869A-16
Perfect score: 1661
Sequence: 1 MDPTISTLDTELTPINGTEE.....EGGGQLPEEILELSGSRLEQ 322

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 484.5 | 29.2 | 324 | 1 | TVRTAS | transforming prote |
| 2 | 462 | 27.8 | 378 | 2 | A39485 | transforming prote |
| 3 | 461.5 | 27.8 | 325 | 1 | TVHUAS | transforming prote |
| 4 | 456.5 | 27.5 | 324 | 2 | S51001 | transforming prote |
| 5 | 418 | 25.2 | 343 | 2 | A35639 | G protein-coupled |
| 6 | 252 | 15.2 | 473 | 2 | JC5835 | anaphylatoxin C3a |
| 7 | 251 | 15.1 | 353 | 2 | C42009 | FMLP-related recep |
| 8 | 250 | 15.1 | 350 | 1 | A37963 | complement C5a ana |
| 9 | 248 | 14.9 | 351 | 2 | B42009 | FMLP-related recep |
| 10 | 243.5 | 14.7 | 369 | 2 | A45291 | somatostatin recep |
| 11 | 242.5 | 14.6 | 355 | 2 | A55733 | G protein-coupled |
| 12 | 241.5 | 14.5 | 369 | 2 | D41795 | somatostatin recep |
| 13 | 236.5 | 14.2 | 346 | 2 | S29248 | somatostatin recep |
| 14 | 236 | 14.2 | 364 | 2 | A49542 | N-formyl peptide c |
| 15 | 230 | 13.8 | 371 | 2 | JC5498 | G protein-coupled |
| 16 | 229 | 13.8 | 352 | 1 | S27357 | complement C5a ana |
| 17 | 226 | 13.6 | 354 | 2 | A23669 | interleukin-8 rece |
| 18 | 226 | 13.6 | 369 | 2 | JC2083 | somatostatin recep |
| 19 | 225 | 13.5 | 352 | 2 | A46520 | N-formyl peptide r |
| 20 | 224.5 | 13.5 | 355 | 2 | JQ1231 | interleukin-8 rece |
| 21 | 224 | 13.5 | 388 | 2 | JN0605 | somatostatin recep |
| 22 | 222.5 | 13.4 | 353 | 2 | JC2492 | G protein-coupled |
| 23 | 221.5 | 13.3 | 363 | 2 | I57940 | somatostatin recep |
| 24 | 221 | 13.3 | 371 | 2 | JC5796 | probable chemoattr |
| 25 | 219.5 | 13.2 | 369 | 2 | B41795 | somatostatin recep |
| 26 | 219.5 | 13.2 | 384 | 2 | A47249 | brain-specific som |
| 27 | 218.5 | 13.2 | 333 | 2 | I38974 | G protein-coupled |
| 28 | 218.5 | 13.2 | 351 | 1 | A46525 | complement C5a ana |
| 29 | 218 | 13.1 | 380 | 2 | S36143 | kappa opioid recep |
| 30 | 217 | 13.1 | 380 | 2 | A48227 | kappa opioid recep |
| 31 | 216.5 | 13.0 | 384 | 2 | JC4629 | somatostatin recep |
| 32 | 215 | 12.9 | 428 | 2 | A44021 | somatostatin recep |
| 33 | 214.5 | 12.9 | 350 | 2 | A42009 | N-formyl peptide r |
| 34 | 214 | 12.9 | 370 | 2 | JC5549 | heptahelical P2Y5- |
| 35 | 214 | 12.9 | 380 | 2 | JC2434 | kappa opioid recep |
| 36 | 213 | 12.8 | 428 | 2 | S30508 | probable G protein |
| 37 | 212.5 | 12.8 | 355 | 2 | JC5067 | G protein-coupled |
| 38 | 212 | 12.8 | 354 | 2 | T09353 | G protein-coupled |
| 39 | 211.5 | 12.7 | 380 | 2 | A55259 | kappa opioid recep |
| 40 | 211 | 12.7 | 391 | 2 | A41795 | somatostatin recep |
| 41 | 211 | 12.7 | 391 | 2 | C41795 | somatostatin recep |
| 42 | 211 | 12.7 | 391 | 2 | A39297 | somatostatin recep |
| 43 | 210.5 | 12.7 | 380 | 2 | JC2338 | kappa opioid recep |
| 44 | 209.5 | 12.6 | 333 | 2 | I65989 | G protein-coupled |
| 45 | 207.5 | 12.5 | 418 | 2 | A46226 | somatostatin recep |

ALIGNMENTS

RESULT 1
TVRTAS
transforming protein mas - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 18-Jun-1999
C;Accession: A31816
R;Young, D.; O'Neill, K.; Jessell, T.; Wigler, M.
Proc. Natl. Acad. Sci. U.S.A. 85, 5339-5342, 1988
A;Title: Characterization of the rat mas oncogene and its high-level expression in the hi
A;Reference number: A31816; MUID:88276953; PMID:2455902
A;Accession: A31816
A;Molecule type: mRNA
A;Residues: 1-324 <YOU>
A;Cross-references: GB:J03823; NID:g205313; PIDN:AAA41573.1; PID:g205314
C;Genetics:
A;Gene: mas
C;Superfamily: mas transforming protein
C;Keywords: G protein-coupled receptor; transforming protein; transmembrane protein
F;31-47/Domain: transmembrane #status predicted <TM1>
F;72-88/Domain: transmembrane #status predicted <TM2>
F;149-165/Domain: transmembrane #status predicted <TM3>
F;185-204/Domain: transmembrane #status predicted <TM4>
F;225-243/Domain: transmembrane #status predicted <TM5>

Query Match 29.2%; Score 484.5; DB 1; Length 324;
Best Local Similarity 39.2%; Pred. No. 1.1e-32;
Matches 112; Conservative 61; Mismatches 86; Indels 27; Gaps 10;

```
Qy         37 VSLVGLTGNAVVLWLLGCRMRRNAFSIYILNLAAADFLFLSGRLI----YSLLSFISIPH 92
              :| :|   | ::|| |  ||||| |::|| :|: ||   |   |    |:|   :|  |
Db         41 ISPLGFVENGILLWFLCFRMRRNPFTVYITHLSIADISLLFCIFILSIDYALDYELSSGH 100

Qy         93 TISKILYPV-MMFSYFAGLSFLSAVSTERCLSVLWPIWYRCHRPTHLSAVVCVLLWALSL 151
                : :   |  :| |  ||  |:|:| |||||||:||||||||| | || || ||||||
Db        101 YYTIVTLSVTFLFGYNTGLYLLTAISVERCLSVLYPIWYRCHRPKHQSAFVCALLWALSC 160

Qy        152 LRSILEWMLCGFLFSGADSAWCQTSD------FITV-AWLIFLCVVLCGSSLVLLIRILC 204
              | : :|:::|  : || :|     ||       || : ::|:|  ::|  || :|:::|
Db        161 LVTTMEYVMC--IDSGEESH--SQSDCRAVIIFIAILSFLVFTPLMLV-SSTILVVKIRK 215

Qy        205 GSRKIPLTRLYVTILLTVLVFLLCGLPFGIQFFLF--LWIHVDREVLFCHVHLVSIFLSA 262
                :     ::||: |::|:::||:  :|  : : |:   |              | ::| :|:  |
Db        216 NTWASHSSKLYIVIMVTIIIFLIFAMPMRVLYLLYYEYW------STFGNLHNISLLFST 269

Qy        263 LNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASEV--DEGGG 306
              :|||||| ||||||| :::: |::||:|| || :|  :    || |
Db        270 INSSANPFIYFFVGSSKKKRFRESLKVVLTRAFKDEMQPRRQEGNG 315
```

RESULT 2
A39485
transforming protein (mrg) - human

```
C;Species: Homo sapiens (man)
C;Date: 28-Feb-1992 #sequence_revision 17-Apr-1993 #text_change 16-Jul-1999
C;Accession: A39485
R;Monnot, C.; Weber, V.; Stinnakre, J.; Bihoreau, C.; Teutsch, B.; Corvol, P.; Clauser,
Mol. Endocrinol. 5, 1477-1487, 1991
A;Title: Cloning and functional characterization of a novel mas-related gene, modulating
A;Reference number: A39485; MUID:92130997; PMID:1723144
A;Accession: A39485
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-378 <MON>
A;Cross-references: GB:S78653; NID:g244209; PIDN:AAB21255.1; PID:g244210
C;Superfamily: mas transforming protein
C;Keywords: G protein-coupled receptor; transmembrane protein

  Query Match             27.8%;  Score 462;  DB 2;  Length 378;
  Best Local Similarity   37.4%;  Pred. No. 9e-31;
  Matches  114;  Conservative   56;  Mismatches  101;  Indels   34;  Gaps    9;

Qy         16 NGTEETL-------CYKQTLSLTVL---TCIVSLVGLTGNAVVLWLLGCRMRRNAFSIYI 65
              | | ||:          :| | | ::     :||| |:  |  | ||| |    | : :||
Db         54 NETNETIHMQMSMAVGQQALPLNIIAPKAVLVSLCGVLLNGTVFWLLCCG-ATNPYMVYI 112

Qy         66 LNLAAADFLFLS----GRLIYSLLSFISIPHTISKILYPVMMFSYFAGLSFLSAVSTERC 121
              |:| ||| ::|     | |  :||::  :   |   |  :  ||:   |  | |:|||||
Db        113 LHLVAADVIYLCCSAVGFLQVTLLTYHGVVFFIPDFLAILSPFSFEVCLCLLVAISTERC 172

Qy        122 LSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQTSD---F 178
              : ||:|||||||| : | ||| |:| |    :|:: :   :         |        |
Db        173 VCVLFPIWYRCHRPKYTSNVVCTLIWGLPFCINIVKSLFLTY--------WKHVKACVIF 224

Qy        179 ITVAWLI--FLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGIQF 236
              : :: |    | :|:| ||| |||| || |::   ||:|  : ::  :|||  ||  :
Db        225 LKLSGLFHAILSLVMCVSSLTLLIRFLCCSQQQKATRVYAVVQISAPMFLLWALPLSVAP 284

Qy        237 FLFLWIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQ 296
               :      | ::     :|:|:||  :|||||||||||||| |::: :::|:::|||||
Db        285 LI-----TDFKMFVTTSYLISLFL-IINSSANPIIYFFVGSLRKKRLKESLRVILQRALA 338

Qy        297 DASEV 301
              |  ||
Db        339 DKPEV 343


RESULT 3
TVHUAS
transforming protein mas - human
C;Species: Homo sapiens (man)
C;Date: 04-Dec-1986 #sequence_revision 04-Dec-1986 #text_change 18-Jun-1999
C;Accession: A01375
R;Young, D.; Waitches, G.; Birchmeier, C.; Fasano, O.; Wigler, M.
Cell 45, 711-719, 1986
A;Title: Isolation and characterization of a new cellular oncogene encoding a protein wi
A;Reference number: A01375; MUID:86218084; PMID:3708691
A;Accession: A01375
A;Molecule type: DNA
A;Residues: 1-325 <YOU>
A;Cross-references: GB:M13150; NID:g187388; PIDN:AAA36199.1; PID:g307158
C;Genetics:
A;Gene: GDB:MAS1
A;Cross-references: GDB:120166; OMIM:165180
A;Map position: 6q24-6q27
C;Superfamily: mas transforming protein
C;Keywords: G protein-coupled receptor; glycoprotein; proto-oncogene; transforming prote
F;31-61/Domain: transmembrane #status predicted <TM1>
F;66-97/Domain: transmembrane #status predicted <TM2>
F;105-135/Domain: transmembrane #status predicted <TM3>
F;150-172/Domain: transmembrane #status predicted <TM4>
F;186-214/Domain: transmembrane #status predicted <TM5>
F;225-250/Domain: transmembrane #status predicted <TM6>
F;258-286/Domain: transmembrane #status predicted <TM7>
F;5,16,22,272/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             27.8%;  Score 461.5;  DB 1;  Length 325;
  Best Local Similarity   38.5%;  Pred. No. 8.7e-31;
  Matches  104;  Conservative   59;  Mismatches   92;  Indels   15;  Gaps    5;

Qy         37 VSLVGLTGNAVVLWLLGCRMRRNAFSIYILNLAAADFLFLSGRLI----YSLLSFISIPH 92
              :| ||   | ::|| |  ||||| |::|| :|: ||   |    |    |:|   :|  |
Db         42 ISPVGFVENGILLWFLCFRMRRNPFTVYITHLSIADISLLFCIFILSIDYALDYELSSGH 101

Qy         93 TISKILYPV-MMFSYFAGLSFLSAVSTERCLSVLWPIWYRCHRPTHLSAVVCVLLWALSL 151
                : :   |  :| |  ||  |:|:| |||||||:||||||||| : ||:|| ||||||
Db        102 YYTIVTLSVTFLFGYNTGLYLLTAISVERCLSVLYPIWYRCHRPKYQSALVCALLWALSC 161

Qy        152 LRSILEWMLCGFLFSGADSAWCQTSDFITVAWLIFLCV--VLCGSSLVLLIRILCGSRKI 209
              | : :|:::|      : |     :  | :| | ||    ::  || :|:::|   :
Db        162 LVTTMEYVMCIDREEESHSRNDCRAVIIFIAILSFLVFTPLMLVSSTILVVKIRKNTWAS 221

Qy        210 PLTRLYVTILLTVLVFLLCGLPFGIQFFLF--LWIHVDREVLFCHVHLVSIFLSALNSSA 267
              ::||: |::|:::||:  :|  : : |:   |         | ::| :|:  | :||||
Db        222 HSSKLYIVIMVTIIIFLIFAMPMRLLYLLYYEYW------STFGNLHHISLLFSTINSSA 275

Qy        268 NPIIYFFVGSFRQRQNRQNLKLVLQRALQD 297
              || ||||||| :::: :::||:|| || :|
Db        276 NPFIYFFVGSSKKKRFKESLKVVLTRAFKD 305


RESULT 4
S51001
transforming protein mas - mouse
N;Alternate names: mas proto-oncogene protein; probable G protein-coupled receptor mas
C;Species: Mus musculus (house mouse)
C;Date: 10-Apr-1996 #sequence_revision 19-Apr-1996 #text_change 20-Apr-2001
C;Accession: S51001; I48647; S29619
R;Metzger, R.; Bader, M.; Ludwig, T.; Berberich, C.; Bunnemann, B.; Ganten, D.
FEBS Lett. 357, 27-32, 1995
A;Title: Expression of the mouse and rat mas proto-oncogene in the brain and peripheral t
A;Reference number: S51001; MUID:95094925; PMID:8001672
A;Accession: S51001
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-324 <MET>
A;Cross-references: EMBL:X67735
R;Metzger, R.; Bader, M.; Ludwig, T.; Berberich, C.; Bunnemann, B.; Ganten, D.
FEBS Lett. 357, 27-32, 1995
A;Title: Expression of the mouse and rat mas proto-oncogene in the brain and peripheral t
A;Reference number: I48647; MUID:95094925; PMID:8001672
A;Accession: I48647
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-87,'I',89-324 <RES>
A;Cross-references: EMBL:X67735; NID:g53011; PIDN:CAA47964.1; PID:g53012
C;Genetics:
A;Gene: mas
C;Superfamily: mas transforming protein
C;Keywords: G protein-coupled receptor; proto-oncogene; transmembrane protein

  Query Match             27.5%;  Score 456.5;  DB 2;  Length 324;
  Best Local Similarity   38.1%;  Pred. No. 2.2e-30;
  Matches  109;  Conservative   61;  Mismatches   89;  Indels   27;  Gaps   10;

Qy         37 VSLVGLTGNAVVLWLLGCRMRRNAFSIYILNLAAADFLFLSGRLI----YSLLSFISIPH 92
              :| :|   | ::|| |  ||||| |::|| :|: ||   |    |  . |:|   :|  |
Db         41 ISPLGFVENGILLWFLCFRMRRNPFTVYITHLSMADISLLFCIFILSTDYALDYELSSGH 100

Qy         93 TISKILYPV-MMFSYFAGLSFLSAVSTERCLSVLWPIWYRCHRPTHLSAVVCVLLWALSL 151
                : :   |  :| |  ||  |:|:| |||||||:||||  ||| | || || || |||
Db        101 HYTIVTLSVTFLFGYNTGLYLLTAISVERCLSVLYPIWYTSHRPKHQSAFVCALLCALSC 160

Qy        152 LRSILEWMLCGFLFSGADSAWCQTSD------FITV-AWLIFLCVVLCGSSLVLLIRILC 204
              | : :|:::|  : || :|     ||       || : ::|:|  ::|  || :|:::|
Db        161 LVTTMEYVMC--IDSGEESH--SRSDCRAVIIFIAILSFLVFTPLMLVSSS-ILVVKIRK 215

Qy        205 GSRKIPLTRLYVTILLTVLVFLLCGLPFGIQFFLF--LWIHVDREVLFCHVHLVSIFLSA 262
                :     ::||: |::|:::||:  :|  : : |:   |         | ::| :|:  |
```

```
Db        216 NTWASHSSKLYIVIMVTIIIFLIFAMPMRVLYLLYYEYW------SAFGNLHNISLLFST 269

Qy        263 LNSSANPIIYFFVGSSFRQRQNRQNLKLVLQRALQDASEV--DEGGG 306
              :|||||| ||||||| :::: |::||:|| || :|  :    || |
Db        270 INSSANPFIYFFVGSSKKKRFRESLKVVLTRAFKDEMQPRRQEGNG 315


RESULT 5
A35639
G protein-coupled receptor RTA - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 16-Jul-1999
C;Accession: A35639
R;Ross, P.C.; Figler, R.A.; Corjay, M.H.; Barber, C.M.; Adam, N.; Harcus, D.R.; Lynch, K
Proc. Natl. Acad. Sci. U.S.A. 87, 3052-3056, 1990
A;Title: RTA, a candidate G protein-coupled receptor: cloning, sequencing, and tissue di
A;Reference number: A35639; MUID:90222168; PMID:2109324
A;Accession: A35639
A;Molecule type: mRNA
A;Residues: 1-343 <ROS>
A;Cross-references: GB:M35297; NID:g206809; PIDN:AAA42087.1; PID:g206810; GB:M32098
C;Superfamily: mas transforming protein
C;Keywords: G protein-coupled receptor; glycoprotein; membrane protein
F;4/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             25.2%;  Score 418;  DB 2;  Length 343;
  Best Local Similarity   35.0%;  Pred. No. 3.6e-27;
  Matches   99;  Conservative   57;  Mismatches  117;  Indels   10;  Gaps    6;

Qy         36 IVSLVGLTGNAVVLWLLGCRMRRNAFSIYILNLAAADFLFLSGRLIYSLL---SFI-SIP 91
              :: | || || :|||  |  ::|  |||| |:||:|| ::|  : : :||   :|: | |
Db         53 LLCLCGLVGNGLVLWFFGFSIKRTPFSIYFLHLASADGIYLFSKAVIALLNMGTFLGSFP 112

Qy         92 HTISKILYPVMMFSYFAGLSFLSAVSTERCLSVLWPIWYRCHRPTHLSAVVCVLLWALSL 151
                : ::   | : ::|||:| | |:| |||:||::|:||   ||  ||| || ||| ||
Db        113 DYVRRVSRIVGLCTFFAGVSLLPAISIERCVSVIFPMWYWRRRPKRLSAGVCALLWLLSF 172

Qy        152 LRSILEWMLCGFLFSGADSAWCQTSDFITVAWLIFL-CVVLCGSSLVLLIRILCGSRKIP 210
              | : :   | ||   |     |   |       | || | ::    | |:|: : | :|:
Db        173 LVTSIHNYFCMFLGHEASGTACLNMDISLGILLFFLFCPLMVLPCLALILHVECRARRRQ 232

Qy        211 LT-RLYVTILLTVLVFLLCGLPFGIQFFLFLWIHVDREVLFCHVHLVSIFLSALNSSANP 269
               : :|   :|  | |||:  :  || :||| |:    ::       |:     :|||| |
Db        233 RSAKLNHVVLAIVSVFLVSSIYLGIDWFLF-WVF---QIPAPFPEYVTDLCICINSSAKP 288

Qy        270 IIYFFVGSFRQRQNRQNLKLVLQRALQDASEVDEGGGQLPEEI 312
              |:||  |  : ::  : |::| ||||:| :|  : .   |  :
Db        289 IVYFLAGRDKSQRLWEPLRVVFQRALRDGAEPGDAASSTPNTV 331


RESULT 6
JC5835
anaphylatoxin C3a receptor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 05-Mar-1998 #sequence_revision 13-Mar-1998 #text_change 17-Mar-2000
C;Accession: JC5835
R;Fukuoka, Y.; Ember, J.A.; Hugli, T.E.
Biochem. Biophys. Res. Commun. 242, 663-668, 1998
A;Title: Cloning and characterization of rat C3a receptor: Differential expression of ra
A;Reference number: JC5835; MUID:98125550; PMID:9464274
A;Accession: JC5835
A;Molecule type: mRNA
A;Residues: 1-473 <FUK>
A;Cross-references: GB:U86379; NID:g3015534; PIDN:AAC40071.1; PID:g2853279
A;Experimental source: brain
C;Comment: This receptor plays a role in the central nervous systems.
C;Superfamily: vertebrate rhodopsin
C;Keywords: glycoprotein; phosphoprotein
F;22-51/Domain: transmembrane #status predicted <TM1>
F;58-83/Domain: transmembrane #status predicted <TM2>
F;96-118/Domain: transmembrane #status predicted <TM3>
F;138-160/Domain: transmembrane #status predicted <TM4>
F;322-347/Domain: transmembrane #status predicted <TM5>
F;367-392/Domain: transmembrane #status predicted <TM6>
F;406-430/Domain: transmembrane #status predicted <TM7>
F;9,201/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;360/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted

  Query Match             15.2%;  Score 252;  DB 2;  Length 473;
  Best Local Similarity   20.5%;  Pred. No. 2.4e-13;
  Matches   95;  Conservative   66;  Mismatches  112;  Indels  190;  Gaps   14;

Qy          7 TLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNAFSIYIL 66
              | ||  | ::      |    | ::  |:  :  |:|| || :|||: | :|:|   ::: |
Db          5 TADTNSTDLH--SRPLFKPQDIASMVILSLTCLLGLPGNGLVLWVAGVKMKRTVNTVWFL 62

Qy         67 NLAAADF-----------LFLSGRLIYSLLSFISIPHTISKILYPVMMFSYFAGLSFLS 114
              :|  |||           | | |   | |         : |::  |:: : || :  |:
Db         63 HLTLADFLCCLSLPFSVAHLILRGHWPYGLF--------LCKLIPSVIILNMFASVFLLT 114

Qy        115 AVSTERCLSVLWPIWYRCHRPTHLSAVVCVLLWALSLL--------RSIL---EWMLCGF 163
              |:| :||| |  ||| : ||    :  ||  :| :: :        | :|   :: :||:
Db        115 AISLDRCLMVHKPIWCQNHRSVRTAFAVCGCVWVVTFVMCIPVFVYRDLLVVDDYSVCGY 174

Qy        164 LFSG-------------------------------------ADSAWCQTS-------- 176
              |                                       |   |  |:
Db        175 NFDSSRAYDYWDYMYNSHLPEINPPDNSTGHVDDRTAPSSSVPARDLWTATTALQSQTFH 234

Qy        177 ------------------------------------------------------------ 176

Db        235 TSPEDPFSQDSASQQPHYGGKPPTVLIATIPGGFPVEDHKSNTLNTGAFLSAHTEPSLTA 294

Qy        177 --------DF-------------------ITVAWLI------FLCVVLCGSSLVLLIRI 202
                      ||                   ||:: |:      |  :: | | :|  :
Db        295 SSSPLYAHDFPDDYFDQLMYGNHAWTPQVAITISRLVVGFLVPFFIMITCYSLIVFRM-- 352

Qy        203 LCGSRKIPLTR-----LYVTILLTVLVFLLCGLPFGIQFFLFLWIHVD---REVLFCHVH 254
                 ||  ||:     | | : : | || :| :|: |   | :   :   |||:    |
Db        353 ----RKTNLTKSRNKTLRVAVAV-VTVFFVCWIPYHIVGILLVITDQESALREVVLPWDH 407

Qy        255 LVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQD 297
              : || |:: ||  || :|  :|   ::: ||::| :|: |  :
Db        408 M-SIALASANSCFNPFLYALLGKDFRKKARQSVKGILEAAFSE 449


RESULT 7
C42009
FMLP-related receptor 2 - human
N;Alternate names: FMLP-related receptor I; probable chemotactic receptor FPRH2
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 26-Aug-1999
C;Accession: C42009
R;Bao, L.; Gerard, N.P.; Eddy Jr., R.L.; Shows, T.B.; Gerard, C.
Genomics 13, 437-440, 1992
A;Title: Mapping of genes for the human C5a receptor (C5AR), human FMLP receptor (FPR), a
A;Reference number: A42009; MUID:92307681; PMID:1612600
A;Accession: C42009
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-353 <BAO>
A;Cross-references: GB:M76673; NID:g182668; PID:g182669
C;Comment: This fMet-Leu-Phe receptor homolog, whose ligand is not yet known, appears not
C;Genetics:
A;Gene: GDB:FPRL2
A;Cross-references: GDB:128855; OMIM:136539
A;Map position: 19q13.3-19q13.4
A;Introns: #status absent
C;Superfamily: vertebrate rhodopsin
C;Keywords: chemotaxis; G protein-coupled receptor; glycoprotein; transmembrane protein

  Query Match             15.1%;  Score 251;  DB 2;  Length 353;
  Best Local Similarity   27.0%;  Pred. No. 2.2e-13;
  Matches   90;  Conservative   61;  Mismatches  126;  Indels   56;  Gaps   13;

Qy         14 PINGTEETLCYKQTLSLTVLTCIVSLV-------GLTGNAVVLWLLGCRMRRNAFSIYIL 66
              |:| ||| |   :    |||     ||       |: || :|:|: | || |   :|  |
```

```
Db          8 PLNETEEVL--PEPAGHTVLWIFSLLVHGVTFVFGVLGNGLVIWVAGFRMTRTVNTICYL 65

Qy         67 NLAAADFLFLSGRLIYSLLS--------FISIPHTISKILYPVMMFSYFAGLSFLSAVST 118

Db         66 NLALADFSF-SAILPFRMVSVAMREKWPFASF---LCKLVHVMIDINLFVSVYLITIIAL 121

Qy        119 ERCLSVLWPIWYRCHRPTHLSAVVCVLLWALSL---LRSILEW-----------MLCGFL 164

Db        122 DRCICVLHPAWAQNHRTMSLAKRVMTGLWIFTIVLTLPNFIFWTTISTTNGDTYCIFNFA 181

Qy        165 FSGADSAWCQTSDFITVAWLIFLCVVLCGSSLVLLIRILC---------GSRKIPLTRLY 215

Db        182 FWG-DTAVERLNVFITMAKVFLILHFIIGFTVPMSIITVCYGIIAAKIHRNHMIKSSRPL 240

Qy        216 VTILLTVLVFLLCGLPFG-IQFFLFLWIHVDREVLF---CHVHLVSI----FLSALNSSA 267

Db        241 RVFAAVVASFFICWFPYELIGILMAVWL---KEMLLNGKYKIILVLINPTSSLAFFNSCL 297

Qy        268 NPIIYFFVGSFRQRQNRQNLKLVLQRALQDASE 300

Db        298 NPILYVFMGRNFQERLIRSLPTSLERALTEVPD 330


RESULT 8
A37963
complement C5a anaphylatoxin receptor - human
C;Species: Homo sapiens (man)
C;Date: 22-Jan-1993 #sequence_revision 14-Feb-1997 #text_change 21-Jul-2000
C;Accession: A37963; S13646; I52417; S30518
R;Boulay, F.; Mery, L.; Tardif, M.; Brouchon, L.; Vignais, P.
Biochemistry 30, 2993-2999, 1991
A;Title: Expression cloning of a receptor for C5a anaphylatoxin on differentiated HL-60
A;Reference number: A37963; MUID:91175748; PMID:2007135
A;Accession: A37963
A;Molecule type: mRNA
A;Residues: 1-350 <BOU>
A;Cross-references: GB:J05327; NID:g179699; PIDN:AAA62831.1; PID:g179700
R;Gerard, N.P.; Gerard, C.
Nature 349, 614-617, 1991
A;Title: The chemotactic receptor for human C5a anaphylatoxin.
A;Reference number: S13646; MUID:91156029; PMID:1847994
A;Accession: S13646
A;Molecule type: mRNA
A;Residues: 1-350 <GER>
A;Cross-references: EMBL:X58674; NID:g29568; PIDN:CAB37830.1; PID:g4467832
R;Gerard, N.P.; Bao, L.; Xiao-Ping, H.; Eddy, R.L.
Biochemistry 32, 1243-1250, 1993
A;Title: Human chemotaxis receptor genes cluster at 19q13.3-13.4. Characterization of th
A;Reference number: I52417; MUID:93192225; PMID:8383526
A;Accession: I52417
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-3 <RES>
A;Cross-references: GB:S56556; GB:S56557; NID:g298577; NID:g298578
C;Genetics:
A;Gene: GDB:C5R1; C5A; C5AR
A;Cross-references: GDB:128856; OMIM:113995
A;Map position: 19q13.3-19q13.4
A;Introns: 1/3
A;Note: the list of introns may be incomplete
C;Function:
A;Description: mediates the inflammatory and chemotactic responses of polymorphonuclear
n
C;Superfamily: vertebrate rhodopsin
C;Keywords: chemotaxis; G protein-coupled receptor; glycoprotein; inflammation; polymorp
F;1-37/Domain: extracellular #status predicted <EX1>
F;38-61/Domain: transmembrane #status predicted <TM1>
F;62-71/Domain: intracellular #status predicted <IN1>
F;72-94/Domain: transmembrane #status predicted <TM2>
F;95-110/Domain: extracellular #status predicted <EX2>
F;111-132/Domain: transmembrane #status predicted <TM3>
F;133-149/Domain: intracellular #status predicted <IN2>
F;150-174/Domain: transmembrane #status predicted <TM4>
F;175-206/Domain: extracellular #status predicted <EX3>
F;207-227/Domain: transmembrane #status predicted <TM5>
F;228-242/Domain: intracellular #status predicted <IN3>
F;243-264/Domain: transmembrane #status predicted <TM6>
F;265-283/Domain: extracellular #status predicted <EX4>
F;284-307/Domain: transmembrane #status predicted <TM7>
F;308-350/Domain: intracellular #status predicted <IN4>
F;5/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match           15.1%;  Score 250;  DB 1;  Length 350;
  Best Local Similarity  26.6%;  Pred. No. 2.6e-13;
  Matches   89;  Conservative   56;  Mismatches  130;  Indels   60;  Gaps   10;

Qy          7 TLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNAFSIYIL 66

Db         19 TLDLN-TPVDKTSNTLRVPDILAL-VIFAVVFLVGVLGNALVVWVTAFEAKRTINAIWFL 76

Qy         67 NLAAADFL-FLSGRLIYSLLSFISIPH-----TISKILYPVMMFSYFAGLSFLSAVSTER 120

Db         77 NLAVADFLSCLALPILFT--SIVQHHHWPFGGAACSILPSLILLNMYASILLLATISADR 134

Qy        121 CLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQTSDFIT 180

Db        135 FLLVFKPIWCQNFRGAGLAWIACAVAWGLALLLTIPSFL-----------------YRV 176

Qy        181 VAWLIFLCVVLCG---------SSLVLLIRILCGS--RKIPLTRLYVTILL--------- 220

Db        177 VREEYFPPKVLCGVDYSHDKRRERAVAIVRLVLGFLWPLLTLTICYTFILLRTWSRRATR 236

Qy        221 ----------TVLVFLLCGLPFGIQFFLFLWIHVDREVLFCHVHLVSIFLS--ALNSSAN 268

Db        237 STKTLKVVVAVVASFFIFWLPYQVTGIMMSFLEPSSPTFLLLNKLDSLCVSFAYINCCIN 296

Qy        269 PIIYFFVGSFRQRQNRQNLKLVLQRALQDASEVDE 303

Db        297 PIIYVVAGQGFQGRLRKSLPSLLRNVLTEESVVRE 331


RESULT 9
B42009
FMLP-related receptor 1 - human
N;Alternate names: FMLP receptor homolog FPR2; formyl peptide receptor like-1; probable c
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 14-Jul-1995 #text_change 21-Jul-2000
C;Accession: B42009; JC1258; JQ1521; A42492; I54751; S21581
R;Bao, L.; Gerard, N.P.; Eddy Jr., R.L.; Shows, T.B.; Gerard, C.
Genomics 13, 437-440, 1992
A;Title: Mapping of genes for the human C5a receptor (C5AR), human FMLP receptor (FPR), a
A;Reference number: A42009; MUID:92307681; PMID:1612600
A;Accession: B42009
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-263,'A',265-338,'C',340-351 <BAO>
A;Cross-references: GB:M76672
A;Note: authors translated the codons GTG for residue 15 as Glu, TCT for residue 19 as Th
R;Perez, H.D.; Holmes, R.; Kelly, E.; McClary, J.; Andrews, W.H.
Gene 118, 303-304, 1992
A;Title: Cloning of a cDNA encoding a receptor related to the formyl peptide receptor of
A;Reference number: JC1258; MUID:92380523; PMID:1511907
A;Accession: JC1258
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-351 <PER>
A;Cross-references: EMBL:X63819; NID:g31460; PID:g31461
A;Experimental source: bone marrow mRNA
R;Ye, R.D.; Cavanagh, S.L.; Quehenberger, O.; Prossnitz, E.R.; Cochrane, C.G.
Biochem. Biophys. Res. Commun. 184, 582-589, 1992
A;Title: Isolation of a cDNA that encodes a novel granulocyte N-formyl peptide receptor
A;Reference number: JQ1521; MUID:92246937; PMID:1374236
A;Accession: JQ1521
A;Molecule type: mRNA
A;Residues: 1-351 <YE2>
A;Cross-references: GB:M88107; NID:g189862; PID:g189863
```

```
A;Experimental source: granulocytes
A;Note: formyl peptide-stimulated calcium mobilization comparable to that of the formyl
R;Murphy, P.M.; Ozcelik, T.; Kenney, R.T.; Tiffany, H.L.; McDermott, D.; Francke, U.
J. Biol. Chem. 267, 7637-7643, 1992
A;Title: A structural homologue of the N-formyl peptide receptor. Characterization and c
A;Reference number: A42492; MUID:92218423; PMID:1373134
A;Accession: A42492
A;Molecule type: mRNA
A;Residues: 1-351 <MUR>
A;Cross-references: GB:M84562; NID:g182741; PIDN:AAA52473.1; PID:g182742
A;Note: sequence extracted from NCBI backbone (NCBIN:94159, NCBIP:94160)
R;Nomura, H.; Nielsen, B.W.; Matsushima, K.
Int. Immunol. 5, 1239-1249, 1993
A;Title: Molecular cloning of cDNAs encoding a LD78 receptor and putative leukocyte chem
A;Reference number: I54751; MUID:94092629; PMID:7505609
A;Accession: I54751
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-351 <RES>
A;Cross-references: GB:D10922; NID:g219864; PIDN:BAA01720.1; PID:g219865
C;Comment: This G-protein coupled receptor, homologous to the N-formyl peptide receptor
differentiated myeloid cells and is probably a chemotactic receptor for some other ligan
C;Genetics:
A;Gene: GDB:FPRL1
A;Cross-references: GDB:127554; OMIM:136538
A;Map position: 19q13.3-19q13.4
A;Introns: #status absent
C;Superfamily: vertebrate rhodopsin
C;Keywords: chemotaxis; G protein-coupled receptor; glycoprotein; transmembrane protein
F;27-53/Domain: transmembrane #status predicted <TM1>
F;59-83/Domain: transmembrane #status predicted <TM2>
F;100-121/Domain: transmembrane #status predicted <TM3>
F;145-169/Domain: transmembrane #status predicted <TM4>
F;206-226/Domain: transmembrane #status predicted <TM5>
F;242-266/Domain: transmembrane #status predicted <TM6>
F;282-307/Domain: transmembrane #status predicted <TM7>
F;4/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;98-176/Disulfide bonds: #status predicted

  Query Match           14.9%;  Score 248;  DB 2;  Length 351;
  Best Local Similarity  27.4%;  Pred. No. 3.9e-13;
  Matches   92;  Conservative   61;  Mismatches  121;  Indels   62;  Gaps   16;

Qy         13 TPINGTEETLCYK-------QTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNAFSIYI 65
              ||:|  || : |:       : | | ||  :  ::|: || :|:|: | || |   :|
Db          7 TPLNEYEE-VSYESAGYTVLRILPLVVLG-VTFVLGVLGNGLVIWVAGFRMTRTVTTICY 64

Qy         66 LNLAAADFLFLSGR--LIYSLLSFISIPH--TISKILYPVMMFSYFAGLSFLSAVSTERC 121
              |||| ||| | :    || |:      |    : |::: |:  : | :  :  :: :||
Db         65 LNLALADFSFTATLPFLIVSMAMGEKWPFGWFLCKLIHIVVDINLFGSVFLIGFIALDRC 124

Qy        122 LSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWML-----------CGFLFSGAD 169
              : || |:| : ||   |:  | |  | |:|: ::  ::             | | |
Db        125 ICVLHPVWAQNHRTVSLAMKVIVGPWILALVLTLPVFLFLTTVTIPNGDTYCTFNF---- 180

Qy        170 SAWCQTSD--------FITVAWLIFLCVVLCGSSLVLLIRILC---------GSRKIPLT 212
              ::|  | :         :|   :|   :   | || : |  :|           |  :
Db        181 ASWGGTPEERLKVAITMLTARGIIRFVI---GFSLPMSIVAICYGLIAAKIHKKGMIKSS 237

Qy        213 RLYVTILLTVLVFLLCGLPFGIQFFL-FLWIHVDREVLFCHVH-LVSIF------LSALN 264
              |    :   | | :|  || :   | :|:   :|:||   : :: |       |:  |
Db        238 RPLRVLTAVVASFFICWFPFQLVALLGTVWL---KEMLFYGKYKIIDILVNPTSSLAFFN 294

Qy        265 SSANPIIYFFVG-SFRQRQNRQNLKLVLQRALQDAS 299
              |  ||::| |||  ||:|    :|   |:||| : |
Db        295 SCLNPMLYVFVGQDFRERL-IHSLPTSLERALSEDS 329


RESULT 10
A45291
somatostatin receptor, somatotropin release-inhibiting factor receptor, SRIF receptor -
C;Species: Rattus norvegicus (Norway rat)
C;Date: 25-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000
C;Accession: A45291
R;Kluxen, F.W.; Bruns, C.; Lubbert, H.
Proc. Natl. Acad. Sci. U.S.A. 89, 4618-4622, 1992
A;Title: Expression cloning of a rat brain somatostatin receptor cDNA.
A;Reference number: A45291; MUID:92262491; PMID:1374909
A;Accession: A45291
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-369 <KLU>
A;Cross-references: GB:M93273; NID:g207026; PIDN:AAA42165.1; PID:g207027
A;Note: sequence extracted from NCBI backbone (NCBIN:102315, NCBIP:102316)
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; transmembrane protein

  Query Match           14.7%;  Score 243.5;  DB 2;  Length 369;
  Best Local Similarity  27.6%;  Pred. No. 9.5e-13;
  Matches   96;  Conservative   57;  Mismatches  126;  Indels   69;  Gaps   13;

Qy          3 PTISTLDTELTPINGTEETLCYKQTLSLTVLTCI---VSLVGLTGNAVVLW-LLGCRMRR 58
              |:   |:  | | ||: :|  |    |  ||| |   | :||| || :|:: :|     :
Db         16 PSPFDLNGSLGPSNGSNQTEPYYDMTSNAVLTFIYFVVCVVGLCGNTLVIYVILRYAKMK 75

Qy         59 NAFSIYILNLAAADFLFLSGRLIYSLLSFISIPH-----TISKILYPVMMFSYFAGLSFL 113
                :|||||| || ||: |  :  |   ::: |       | :::  |   : |  :  |
Db         76 TITNIYILNLAIADELFMLG--LPFLAMQVALVHWPFGKAICRVVMTVDGINQFTSIFCL 133

Qy        114 SAVSTERCLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWML------------ 160
              : :| :| |:|: ||      |    : :: | :| :|||  ||  |:
Db        134 TVMSIDRYLAVVHPIKSAKWRRPRTAKMINVAVWGVSLL-VILPIMIYAGLRSNQWGRSS 192

Qy        161 CGFLFSGADSAWCQTSDFITVAWLIFLCV---VLCGSSLVLLIR-----ILCGSRKIPLT 212
              |   : |   ||   : ||  |:::   |   ::|   | ::|:     |  || |   :
Db        193 CTINWPGESGAW--YTGFIIYAFILGFLVPLTIICLCYLFIIIKVKSSGIRVGSSKRKKS 250

Qy        213 RLYVT--ILLTVLVFLLCGLPFGIQFFLFLWIHVDREVLFCHVHLVS------------ 257
              ||  : : | ||: | ||    |::|           :|  ||
Db        251 EKKVTRMVSIVVAVFIFCWLP----FYIF-----------NVSSVSVAISPTPALKGMF 294

Qy        258 ---IFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASEVD 302
                : |:  || ||||:| |:    ::  :  | ||      :|    |
Db        295 DFVVILTYANSCANPILYAFLSDNFKKSFQNVLCLVKVSGAEDGERSD 342


RESULT 11
A55733
G protein-coupled receptor GPR1 - human
C;Species: Homo sapiens (man)
C;Date: 07-Jul-1995 #sequence_revision 07-Jul-1995 #text_change 21-Jul-2000
C;Accession: A55733
R;Marchese, A.; Docherty, J.M.; Nguyen, T.; Heiber, M.; Cheng, R.; Heng, H.H.Q.; Tsui, L.
Genomics 23, 609-618, 1994
A;Title: Cloning of human genes encoding novel G protein-coupled receptors.
A;Reference number: A55733; MUID:95154831; PMID:7851889
A;Accession: A55733
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-355 <MAR>
A;Cross-references: GB:U13666; NID:g577412; PIDN:AAA64592.1; PID:g577413
C;Genetics:
A;Gene: GDB:GPR1
A;Cross-references: GDB:371707; OMIM:600239
A;Map position: 15q25-15q26.1
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor

  Query Match           14.6%;  Score 242.5;  DB 2;  Length 355;
  Best Local Similarity  23.0%;  Pred. No. 1.1e-12;
  Matches   75;  Conservative   59;  Mismatches   97;  Indels   95;  Gaps   10;

Qy         32 VLTCIVSLVGLTGNAVVLWLLGCRMRRNAFSIYILNLAAADFLFLSGRLIYSLLSFISIP 91
              || |:  ::|: |||:|:|  | : ::   ::: |||| |||:||          : :|
Db         43 VLYCLAFVLGIPGNAIVIWFTGLKWKKTVTTLWFLNLAIADFIFL----------LFLP 91
```

```
Qy          92 HTISKILYPVMMFSY-----------------FAGLSFLSAVSTERCLSVLWPIWYRCH 133
               ||   |  | | :                  || : ||: :| :  : :: |:    |
Db          92 LYIS---YVAMNFHWPFGIWLCKANSFTAQLNMFASVFFLTVISLDHYIHLIHPVLSHRH 148

Qy         134 RPTHLSAVVCVLLWALSLL---------------------------------RSILEW 158
               |    | :| : :| |: |                                 :| |
Db         149 RTLKNSLIVIIFIWLLASLIGGPALYFRDTVEFNNHTLCYNNFQKHDPDLTLIRHHVLTW 208

Qy         159 M--LCGFLFSGADSAWCQTSDFITVAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYV 216
               :  : |:||            :|:: : :||::       :||           :| :
Db         209 VKFIIGYLF-----------PLLTMS-ICYLCLIFKVKKRTVLIS----------SRHFW 246

Qy         217 TILLTVLVFLLCGLPFGIQFFLFLWIH---VDREVLFCHVHLVSIFLSALNSSANPIIYF 273
               |||: |: |::|  |: :     | ||       |:   : | |  |: |||  |||:|
Db         247 TILVVVVAFVVCWTPYHLFSIWELTIHHNSYSHHVMQAGIPL-STGLAFLNSCLNPILYV 305

Qy         274 FVGSFRQRQNRQNLKLVLQRALQDAS 299
               :    | : | ::  :|:  | : |
Db         306 LISKKFQARFRSSVAEILKYTLWEVS 331
```

RESULT 12
D41795
somatostatin receptor 2 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 24-Nov-1999
C;Accession: D41795; I56236
R;Yamada, Y.; Post, S.R.; Wang, K.; Tager, H.S.; Bell, G.I.; Seino, S.
Proc. Natl. Acad. Sci. U.S.A. 89, 251-255, 1992
A;Title: Cloning and functional characterization of a family of human and mouse somatost
A;Reference number: A41795; MUID:92108031; PMID:1346068
A;Accession: D41795
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-369 <YAM>
A;Cross-references: GB:M81832; NID:g201060; PIDN:AAA58256.1; PID:g201061
R;Elliott, D.E.; Metwali, A.; Blum, A.M.; Sandor, M.; Lynch, R.; Weinstock, J.V.
J. Immunol. 153, 1180-1186, 1994
A;Title: T lymphocytes isolated from the hepatic granulomas of schistosome-infected mice
A;Reference number: I56236; MUID:94300079; PMID:7913111
A;Accession: I56236
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 99-309 <RES>
A;Cross-references: GB:S71756; NID:g560631
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; hormone receptor; transmembrane protein

  Query Match  14.5%;  Score 241.5;  DB 2;  Length 369;
  Best Local Similarity  27.7%;  Pred. No. 1.4e-12;
  Matches  95;  Conservative  56;  Mismatches  123;  Indels  69;  Gaps  13;

```
Qy           8 LDTELTPINGTEETLCYKQTLSLTVLTCI---VSLVGLTGNAVVLW-LLGCRMRRNAFSI 63
               |:  | | ||: :|  |     |  ||| |   | :||| || :|:: :|     :   :|
Db          21 LNGSLGPSNGSNQTEPYYDMTSNAVLTFIYFVVCVVGLCGNTLVIYVILRYAKMKTITNI 80

Qy          64 YILNLAAADFLFLSGRLIYSLLSFISIPH-----TISKILYPVMMFSYFAGLSFLSAVST 118
               |||||| || ||: |  :  |   ::: |      | :::  |   : |  :|: :|
Db          81 YILNLAIADELFMLG--LPFLAMQVALVHWPFGKAICRVVMTVDGINQFTSIFCLTVMSI 138

Qy         119 ERCLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWML-------------CGFLF 165
               :| |:|: ||      |    : :: | :| :|||  ||  |:                |   :
Db         139 DRYLAVVHPIKSAKWRRPRTAKMINVAVWCVSLL-VILPIMIYAGLRSNQWGRSSCTINW 197

Qy         166 SGADSAWCQTSDFITVAWLIFLCV---VLCGSSLVLLIR-----ILCGSRKIPLTRLYVT 217
                |    ||   : ||  |:::   |   ::|    | ::|:      |  || |   :   || .
Db         198 PGESGAW--YTGFIIYAFILGFLVPLTIICLCYLFIIIKVKSSGIRVGSSKRKKSEKKVT 255

Qy         218 --ILLTVLVFLLCGLPFGIQFFLFLWIHVDREVLFCHVHLVS----------------IF 259
                 : : | ||: | ||    |::|            :|  ||                :
Db         256 RMVSIVVAVFIFCWLP----FYIF-----------NVSSVSVAISPTPALKGMFDFVVI 299

Qy         260 LSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASEVD 302
```

```
               |:  || ||||:| |:    ::  :  | ||     :|    |
Db         300 LTYANSCANPILYAFLSDNFKKSFQNVLCLVKVSGTEDGERSD 342
```

RESULT 13
S29248
somatostatin receptor 2B - mouse
C;Species: Mus musculus (house mouse)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 24-Nov-1999
C;Accession: S29248
R;Vanetti, M.; Kouba, M.; Wang, X.; Vogt, G.; Hoellt, V.
FEBS Lett. 311, 290-294, 1992
A;Title: Cloning and expression of a novel mouse somatostatin receptor (SSTR2B).
A;Reference number: S29248; MUID:93012001; PMID:1397330
A;Accession: S29248
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-346 <VAN>
A;Cross-references: EMBL:X68951; NID:g54197; PIDN:CAA48766.1; PID:g54198
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; transmembrane protein

  Query Match  14.2%;  Score 236.5;  DB 2;  Length 346;
  Best Local Similarity  27.9%;  Pred. No. 3.4e-12;
  Matches  95;  Conservative  56;  Mismatches  121;  Indels  69;  Gaps  13;

```
Qy           8 LDTELTPINGTEETLCYKQTLSLTVLTCI---VSLVGLTGNAVVLW-LLGCRMRRNAFSI 63
               |:  | | ||: :|  |     |  ||| |   | :||| || :|:: :|     :   :|
Db          21 LNGSLGPSNGSNQTEPYYDMTSNAVLTFIYFVVCVVGLCGNTLVIYVILRYAKMKTITNI 80

Qy          64 YILNLAAADFLFLSGRLIYSLLSFISIPH-----TISKILYPVMMFSYFAGLSFLSAVST 118
               |||||| || ||: |  :  |   ::: |      | :::  |   : |  :|: :|
Db          81 YILNLAIADELFMLG--LPFLAMQVALVHWPFGKAICRVVMTVDGINQFTSIFCLTVMSI 138

Qy         119 ERCLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWML-------------CGFLF 165
               :| |:|: ||      |    : :: | :| :|||  ||  |:                |   :
Db         139 DRYLAVVHPIKSAKWRRPRTAKMINVAVWCVSLL-VILPIMLYAGLRSNQWGRSSCTINW 197

Qy         166 SGADSAWCQTSDFITVAWLIFLCV---VLCGSSLVLLIR-----ILCGSRKIPLTRLYVT 217
                |    ||   : ||  |:::   |   ::|    | ::|:      |  || |   :   || .
Db         198 PGESGAW--YTGFIIYAFILGFLVPLTIICLCYLFIIIKVKSSGIRVGSSKRKKSEKKVT 255

Qy         218 --ILLTVLVFLLCGLPFGIQFFLFLWIHVDREVLFCHVHLVS----------------IF 259
                 : : | ||: | ||    |::|            :|  ||                :
Db         256 RMVSIVVAVFIFCWLP----FYIF-----------NVSSVSVAISPTPALKGMFDFVVI 299

Qy         260 LSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASE 300
               |:  |: ||||:| |:    ::  :  | ||     | :|
Db         300 LTYANTCANPILYAFLSDNFKKSFQNVLCLVKADNSQSGAE 340
```

RESULT 14
A49542
N-formyl peptide chemotactic receptor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 23-Mar-1995 #sequence_revision 23-Mar-1995 #text_change 05-Nov-1999
C;Accession: A49542
R;Gao, J.L.; Murphy, P.M.
J. Biol. Chem. 268, 25395-25401, 1993
A;Title: Species and subtype variants of the N-formyl peptide chemotactic receptor reveal
A;Reference number: A49542; MUID:94064602; PMID:8244972
A;Accession: A49542
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-364 <GAO>
A;Cross-references: GB:L22181; NID:g347396; PIDN:AAA16110.1; PID:g347397
C;Genetics:
A;Introns: #status absent
C;Superfamily: vertebrate rhodopsin
C;Keywords: chemotaxis; transmembrane protein

  Query Match  14.2%;  Score 236;  DB 2;  Length 364;
  Best Local Similarity  24.8%;  Pred. No. 3.9e-12;

```
  Matches   84;  Conservative   57;  Mismatches  132;  Indels   66;  Gaps   13;

Qy         32 VLTCIVSLVGLTGNAVVLWLLGCRMRRNAFSIYILNLAAADFLFLSGRLIYSLLSFISIP 91
              ::  :  .::|: || :|:|: | ||:    :|  |||| ||| | |  | : : | :
Db         39 LIFAVTFVLGVLGNGLVIWVAGFRMKHTVTTISYLNLAIADFCFTS-TLPFYIASMVMGG 97

Qy         92 H-----TISKILYPVMMFSYFAGLSFLSAVSTERCLSVLWPIWYRCHRPTHLSAVVCVLL 146
              |      : | :| |:  : |  :  :: :: :||: || |:| : ||   |:  | ::
Db         98 HWPFGWFMCKFIYTVIDINLFGSVFLIALIALDRCICVLHPVWAQNHRTVSLAKKVIIVP 157

Qy        147 WALSLLRSILEWMLCGFLFS---------------GADSAWCQTSDF-----------I 179
              |             :| || :                |.    | | ||          :
Db        158 W------------ICAFLLTLPVIIRLTTVPNSRLGPGKTAC-TFDFSPWTKDPVEKRKV 204

Qy        180 TVAWLIFLCVV--LCGSSLVLLIRILC---------GSRKIPLTRLYVTILLTVLVFLLC 228
               |  |    ::  : | |  : |  :|             |  :|    :   |  | ||
Db        205 AVTMLTVRGIIRFIIGFSTPMSIVAICYGLITTKIHRQGLIKSSRPLRVLSFVVAAFFLC 264

Qy        229 GLPFGIQFFLFLWIHVDREVLFCHVHLVSIF-----LSALNSSANPIIYFFVG-SFRQRQ 282
               || :  |   | |  :      :|:       |:  ||  ||::| |:|  ||:|
Db        265 WCPFQV-VALISTIQVRERLKNMTPGIVTALKITSPLAFFNSCLNPMLYVFMGQDFRERL 323

Qy        283 NRQNLKLVLQRAL-QDASEVDEGGGQLPEEILELSGSRL 320
              :|   |:||| :|:::  : | |      || : |
Db        324 -IHSLPASLERALTEDSAQTSDTGTNLGTNSTSLSENTL 361


RESULT 15
JC5498
G protein-coupled receptor DEZ - mouse
C;Species: Mus musculus (house mouse)
C;Date: 07-Jul-1997 #sequence_revision 29-Aug-1997 #text_change 19-May-2000
C;Accession: JC5498
R;Methner, A.; Hermey, G.; Schinke, B.; Hermans-Borgmeyer, I.
Biochem. Biophys. Res. Commun. 233, 336-342, 1997
A;Title: A novel G protein-coupled receptor with homology to neuropeptide and chemoattra
A;Reference number: JC5498; MUID:97289630; PMID:9144535
A;Contents: Brain
A;Accession: JC5498
A;Molecule type: mRNA
A;Residues: 1-371 <MET>
A;Cross-references: GB:U79525; NID:g1732346; PIDN:AAB53789.1; PID:g1732347
C;Comment: This protein is involved in the bone metabolism.
C;Superfamily: vertebrate rhodopsin
F;110-187/Disulfide bonds: #status predicted

  Query Match             13.8%;  Score 230;  DB 2;  Length 371;
  Best Local Similarity   25.2%;  Pred. No. 1.2e-11;
  Matches   82;  Conservative   57;  Mismatches  122;  Indels   64;  Gaps   12;

Qy         30 LTVLTCIVSLVGLTGNAVVLWLLGCRMRRNAFSIYILNLAAADFL---FLSGRLIYSLLS 86
              | |:  :|  :|| || :|: :   :|::   ::: :||| ||||   ||   : |: :
Db         41 LVVIYSLVCFLGLLGNGLVIVIATFKMKKTVNTVWFVNLAVADFLFNIFLPMHITYAAMD 100

Qy         87 FISI-PHTISKILYPVMMFSYFAGLSFLSAVSTERCLSVLWPIWYRCHRPTHLSAVVCVL 145
              :  :    : ||   ::  : :  :  |: :| :||:||| |:| : ||   |: : |
Db        101 YHWVFGKAMCKISNFLLSHNMYTSVFLLTVISFDRCISVLLPVWSQNHRSIRLAYMTCSA 160

Qy        146 LWALSLLRSILEWMLCGFLFSGADSAWCQTSDF---ITVAWLIFLC----------VV 190
              :| |:             |  |     :  |::    ||      |            ||
Db        161 VWVLA------------FFLSSPSLVFRDTANIHGKITCFNNFSLAAPESSPHPAHSQVV 208

Qy        191 LCGSS---LVLLIRILCGSRKIP---LTRLYVT-----------------ILLTVLV-F 225
              | |     | : | |||   ||    :|  |:|      .           |::|::: |
Db        209 STGYSRHVAVTVTRFLCGF-LIPVFIITACYLTIVFKLQRNRLAKNKKPFKIIITIIITF 267

Qy        226 LLCGLPFGIQFFLFLWIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVG-SFRQRQNR 284
              ||  |:   : | |         :|     ::  ::  ||  |||:| |:|  ||:
Db        268 FLCWCPYHTLYLLELHHTAVPSSVFSLGLPLATAVAIANSCMNPILYVFMGHDFRK---- 323

Qy        285 QNLKLVLQRALQDASEVDEGGGQLP 309
               |: |   | :|   | |    |
Db        324 --FKVALFSRLANALSEDTGPSSYP 346
```

```
Search completed: July  2, 2003, 19:22:54
Job time : 21 secs
```

GenCore version 5.1.6


OM protein - protein search, using sw model

Run on: July 2, 2003, 19:13:29 ; Search time 72 Seconds
(without alignments)
595.926 Million cell updates/sec

Title: US-09-849-869A-16
Perfect score: 1661
Sequence: 1 MDPTISTLDTELTPINGTEE.....EGGGQLPEEILELSGSRLEQ 322

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 1661 | 100.0 | 322 | 22 | AAE12794 | Human G protein co |
| 2 | 1661 | 100.0 | 322 | 22 | AAU04371 | Human G-protein co |
| 3 | 1661 | 100.0 | 322 | 22 | AAG64294 | Human GTP-binding |
| 4 | 1661 | 100.0 | 322 | 23 | AAE21288 | Human MrgX1 (mas-r |
| 5 | 1661 | 100.0 | 322 | 23 | AAE17074 | Human G-protein co |
| 6 | 1652 | 99.5 | 322 | 20 | AAY30162 | Human dorsal root |
| 7 | 1624 | 97.8 | 322 | 20 | AAY30161 | Human dorsal root |
| 8 | 1390 | 83.7 | 322 | 21 | AAB14846 | Human novel G-prot |
| 9 | 1383 | 83.3 | 322 | 21 | AAY90761 | Human G protein-co |
| 10 | 1383 | 83.3 | 322 | 23 | AAE21296 | Human MrgX3 (mas-r |
| 11 | 1383 | 83.3 | 337 | 23 | AAU97598 | Human G-protein co |
| 12 | 1378 | 83.0 | 322 | 21 | AAY90762 | Human G protein-co |
| 13 | 1373 | 82.7 | 322 | 20 | AAY30159 | Human dorsal root |
| 14 | 1314 | 79.1 | 322 | 21 | AAY87664 | Human G protein-co |
| 15 | 1313 | 79.0 | 322 | 20 | AAY30160 | Human dorsal root |
| 16 | 1294 | 77.9 | 322 | 20 | AAY30163 | Human dorsal root |
| 17 | 1294 | 77.9 | 322 | 21 | AAY87663 | Human G protein-co |
| 18 | 1294 | 77.9 | 322 | 23 | AAE21297 | Human MrgX4 (mas-r |
| 19 | 1287 | 77.5 | 322 | 20 | AAY30164 | Human dorsal root |
| 20 | 1038.5 | 62.5 | 330 | 22 | AAU04376 | Human G-protein co |
| 21 | 1038.5 | 62.5 | 330 | 22 | AAG62852 | Amino acid sequenc |
| 22 | 1038.5 | 62.5 | 330 | 22 | AAG64301 | Human GTP-binding |
| 23 | 1038.5 | 62.5 | 330 | 22 | AAY72910 | Human IGS3 G-prote |
| 24 | 1038.5 | 62.5 | 330 | 22 | AAB67654 | Amino acid sequenc |
| 25 | 1038.5 | 62.5 | 330 | 23 | AAE21289 | Human MrgX2 (mas-r |
| 26 | 1038.5 | 62.5 | 330 | 23 | ABB06258 | Human G protein-co |
| 27 | 1038.5 | 62.5 | 330 | 23 | ABB08772 | Human G-protein co |
| 28 | 1038.5 | 62.5 | 330 | 23 | AAE15639 | Human G-protein co |
| 29 | 1038.5 | 62.5 | 330 | 23 | AAU10589 | Human G-protein co |
| 30 | 1038.5 | 62.5 | 330 | 23 | AAE17072 | Human G-protein co |
| 31 | 1033.5 | 62.2 | 330 | 23 | AAU10590 | Human G-protein co |
| 32 | 1032 | 62.1 | 330 | 22 | AAE09974 | G-protein coupled- |
| 33 | 984 | 59.2 | 324 | 22 | AAG80957 | Human nGPCR53. Ho |
| 34 | 885 | 53.3 | 264 | 22 | AAM78329 | Human protein SEQ |
| 35 | 847 | 51.0 | 337 | 20 | AAY30158 | Rat dorsal root re |
| 36 | 821 | 49.4 | 338 | 23 | AAE21301 | Mouse MrgB2 (mas-r |
| 37 | 820 | 49.4 | 338 | 22 | AAB73294 | G-protein coupled |
| 38 | 779 | 46.9 | 273 | 23 | AAE21323 | Mouse MrgB10 (mas- |
| 39 | 773.5 | 46.6 | 325 | 23 | ABB08771 | Mouse G-protein co |
| 40 | 769.5 | 46.3 | 338 | 23 | AAE21300 | Mouse MrgB1 (mas-r |
| 41 | 764.5 | 46.0 | 303 | 23 | AAE21307 | Mouse MrgA9 (mas-r |
| 42 | 761.5 | 45.8 | 331 | 22 | AAE12793 | Rat G protein coup |
| 43 | 760.5 | 45.8 | 338 | 21 | AAY77736 | Seven times membra |
| 44 | 756.5 | 45.5 | 304 | 23 | AAE21281 | Mouse MrgA1 (mas-r |
| 45 | 748.5 | 45.1 | 305 | 23 | AAE21309 | Mouse MrgA11 (mas- |

ALIGNMENTS

RESULT 1
AAE12794
ID AAE12794 standard; Protein; 322 AA.
XX
AC AAE12794;
XX
DT 15-JAN-2002 (first entry)
XX
DE Human G protein coupled receptor 1 (GPCR1).
XX
KW Human; G protein coupled receptor; GPCR; purine receptor; vasodilation;
KW therapy; central nervous system disorder; sedative; antiinflammatory;
KW platelet aggregation; gastric secretion; apoptosis; neuroprotective;
KW inflammation; anticonvulsant; anticoagulant; bronchoconstriction;
KW bronchodilator; vasoconstrictor.
XX
OS Homo sapiens.
XX
PN WO200174902-A2.
XX
PD 11-OCT-2001.
XX
PF 02-APR-2001; 2001WO-EP03688.
XX
PR 04-APR-2000; 2000GB-0008252.
XX
PA (JANC ) JANSSEN PHARM NV.
XX
PI Bender E, Buist A, Ercken M, Baggerman GLE, Jurzak M, Schoofs LAH;
PI Luyten WHML;
XX
DR WPI; 2001-648551/74.
DR N-PSDB; AAD20943.

```
XX
PT   New G protein coupled receptor polypeptide, adenine binding G protein
PT   coupled receptor, useful in treatment of diseases in which pathology is
PT   associated with activity at purine receptors e.g. central nervous
PT   system disorders -
XX
PS   Claim 4; Fig 1; 81pp; English.
XX
CC   The invention relates to G protein coupled receptor (adenine binding
CC   GPCR) and its corresponding nucleic acids. GPCR DNA is used to prepare
CC   medicaments for treating diseases in which pathology is associated with
CC   activity at purine receptors (especially mutation or downregulation of
CC   expression of native adenine binding GPCRs). GPCR DNA and its protein
CC   are used to treat central nervous system disorders (e.g. as sedatives,
CC   anticonvulsives, etc.), as stimulators of NO production by vascular
CC   endothelial cells, as inhibitors of platelet aggregation, gastric
CC   secretion or in diseases associated with apoptosis, vasodilation,
CC   bronchoconstriction or inflammation. They are useful to identify
CC   interacting compounds useful to treat these diseases. They are used to
CC   generate antibodies, useful to detect polypeptide (e.g. diagnostically),
CC   in polypeptide purification and therapeutically. GPCR DNA is also used
CC   to produce probes and primers, useful to detect/quantify polynucleotides
CC   (e.g. diagnostically or in tissue distribution studies) or to isolate
CC   similar sequences (e.g. from other species). The present sequence is
CC   human GPCR1 protein.
XX
SQ   Sequence   322 AA;

  Query Match             100.0%;  Score 1661;  DB 22;  Length 322;
  Best Local Similarity   100.0%;  Pred. No. 4e-171;
  Matches  322;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MDPTISTLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDPTISTLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNA 60

Qy         61 FSIYILNLAAADFLFLSGRLIYSLLSFISIPHTISKILYPVMMFSYFAGLSFLSAVSTER 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 FSIYILNLAAADFLFLSGRLIYSLLSFISIPHTISKILYPVMMFSYFAGLSFLSAVSTER 120

Qy        121 CLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQTSDFIT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQTSDFIT 180

Qy        181 VAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGIQFFLFL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGIQFFLFL 240

Qy        241 WIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 WIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASE 300

Qy        301 VDEGGGQLPEEILELSGSRLEQ 322
              |||||||||||||||||||||||
Db        301 VDEGGGQLPEEILELSGSRLEQ 322


RESULT 2
AAU04371
ID   AAU04371 standard; Protein; 322 AA.
XX
AC   AAU04371;
XX
DT   23-OCT-2001  (first entry)
XX
DE   Human G-protein coupled receptor, hRUP17.
XX
KW   Human; G-protein coupled receptor; GPCR; hRUP17; agonist;
KW   inverse agonist; lung cancer.
XX
OS   Homo sapiens.
XX
PN   WO200136471-A2.
XX
PD   25-MAY-2001.
XX
PF   16-NOV-2000; 2000WO-US31509.
XX
PR   17-NOV-1999;   99US-0166088.
PR   17-NOV-1999;   99US-0166099.
PR   17-NOV-1999;   99US-0166369.
PR   23-DEC-1999;   99US-0171900.
PR   23-DEC-1999;   99US-0171901.
PR   23-DEC-1999;   99US-0171902.
PR   11-FEB-2000; 2000US-0181749.
PR   14-MAR-2000; 2000US-0189258.
PR   14-MAR-2000; 2000US-0189259.
PR   10-APR-2000; 2000US-0195898.
PR   10-APR-2000; 2000US-0195899.
PR   10-APR-2000; 2000US-0196078.
PR   28-APR-2000; 2000US-0200419.
PR   12-MAY-2000; 2000US-0203630.
PR   12-JUN-2000; 2000US-0210741.
PR   12-JUN-2000; 2000US-0210982.
PR   21-AUG-2000; 2000US-0226760.
PR   26-SEP-2000; 2000US-0235418.
PR   26-SEP-2000; 2000US-0235779.
PR   20-OCT-2000; 2000US-0242332.
PR   20-OCT-2000; 2000US-0242343.
XX
PA   (AREN-) ARENA PHARM INC.
XX
PI   Chen R,  Dang HT,  Lowitz KP;
XX
DR   WPI; 2001-355616/37.
DR   N-PSDB; AAS07944.
XX
PT   Endogenous and non-endogenous versions of human G-protein coupled
PT   receptors for direct identification of candidate compounds as agonists,
PT   inverse agonists or partial agonists for use as therapeutic agents -
XX
PS   Claim 37; Page 107; 160pp; English.
XX
CC   The sequence represents a human G-protein coupled receptor (GPCR),
CC   hRUP17. The endogenous and non-endogenous, constitutively activated
CC   versions of human G-protein coupled receptors (GPCR), are useful for
CC   direct identification of candidate compounds as receptor agonists,
CC   inverse agonists or partial agonists having applicability as therapeutic
CC   agents for treating diseases related to GPCR, e.g. lung cancer.
CC   Non-endogenous version of human GPCRs are also utilized in research
CC   settings and in vitro and in vivo system, incorporating GPCRs can be
CC   utilised to elucidate and understand the roles these receptors
CC   play in the human condition, both normal and diseased.
XX
SQ   Sequence   322 AA;

  Query Match             100.0%;  Score 1661;  DB 22;  Length 322;
  Best Local Similarity   100.0%;  Pred. No. 4e-171;
  Matches  322;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MDPTISTLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDPTISTLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNA 60

Qy         61 FSIYILNLAAADFLFLSGRLIYSLLSFISIPHTISKILYPVMMFSYFAGLSFLSAVSTER 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 FSIYILNLAAADFLFLSGRLIYSLLSFISIPHTISKILYPVMMFSYFAGLSFLSAVSTER 120

Qy        121 CLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQTSDFIT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQTSDFIT 180

Qy        181 VAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGIQFFLFL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGIQFFLFL 240
```

```
Qy        241 WIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 WIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASE 300

Qy        301 VDEGGGQLPEEILELSGSRLEQ 322
              ||||||||||||||||||||||
Db        301 VDEGGGQLPEEILELSGSRLEQ 322
```

RESULT 3
AAG64294
ID AAG64294 standard; Protein; 322 AA.
XX
AC AAG64294;
XX
DT 21-SEP-2001 (first entry)
XX
DE Human GTP-binding protein-coupled receptor GPRv20.
XX
KW GTP-binding protein-coupled receptor; neuroprotective; immunomodulatory;
KW muscular; urinary; circulatory; anorectic; human; guanosine triphosphate;
KW G-protein.
XX
OS Homo sapiens.
XX
PN WO200148189-A1.
XX
PD 05-JUL-2001.
XX
PF 28-DEC-2000; 2000WO-JP09409.
XX
PR 28-DEC-1999; 99JP-0375152.
PR 31-MAR-2000; 2000JP-0101339.
PR 23-MAY-2000; 2000JP-0155978.
XX
PA (HELI-) HELIX RES INST.
XX
PI Matsumoto S, Oda T, Saito Y, Morikawa N, Yoshida K, Suwa M;
PI Sugiyama T;
XX
DR WPI; 2001-425663/45.
DR N-PSDB; AAH49507.
XX
PT Family of guanosine triphosphate binding protein coupled receptors and
PT genes encoding them for treatment and prevention of diseases associated
PT with these receptors -
XX
PS Claim 1; Pages 75-78; 137pp; Japanese.
XX
CC The present sequence is the protein sequence for a human guanosine
CC triphosphate (GTP)-binding protein-coupled receptor. The receptor is
CC useful for the investigation, diagnosis, treatment and prevention of
CC diseases associated with GTP-binding protein-coupled receptors, including
CC neurological, circulatory, digestive system, immune system, muscle and
CC urinary system disorders. GTP-binding proteins are also known as
CC G-proteins.
XX
SQ Sequence 322 AA;

```
  Query Match             100.0%;  Score 1661;  DB 22;  Length 322;
  Best Local Similarity   100.0%;  Pred. No. 4e-171;
  Matches  322;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MDPTISTLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDPTISTLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNA 60

Qy         61 FSIYILNLAAADFLFLSGRLIYSLLSFISIPHTISKILYPVMMFSYFAGLSFLSAVSTER 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 FSIYILNLAAADFLFLSGRLIYSLLSFISIPHTISKILYPVMMFSYFAGLSFLSAVSTER 120

Qy        121 CLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQTSDFIT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQTSDFIT 180

Qy        181 VAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGIQFFLFL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGIQFFLFL 240

Qy        241 WIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 WIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASE 300

Qy        301 VDEGGGQLPEEILELSGSRLEQ 322
              ||||||||||||||||||||||
Db        301 VDEGGGQLPEEILELSGSRLEQ 322
```

RESULT 4
AAE21288
ID AAE21288 standard; Protein; 322 AA.
XX
AC AAE21288;
XX
DT 01-JUL-2002 (first entry)
XX
DE Human MrgX1 (mas-related gene) protein.
XX
KW Human; mas-related gene; G-protein coupled receptor; drg-12 protein;
KW receptor; sensory perception; pain; analgesic; MrgX1.
XX
OS Homo sapiens.
XX
PN WO200183555-A2.
XX
PD 08-NOV-2001.
XX
PF 04-MAY-2001; 2001WO-US14519.
XX
PR 04-MAY-2000; 2000US-202027P.
PR 01-AUG-2000; 2000US-222344P.
PR 03-NOV-2000; 2000US-0704707.
PR 19-APR-2001; 2001US-285493P.
XX
PA (CALY ) CALIFORNIA INST OF TECHNOLOGY.
XX
PI Anderson DJ, Dong X, Zylka M, Han S, Simon M;
XX
DR WPI; 2002-171346/22.
DR N-PSDB; AAD33744.
XX
PT Isolated polypeptide, Mrg, which is a G-protein coupled receptor and an
PT isolated polypeptide, drg-12, which is also a receptor, useful for
PT identifying agonists or antagonists for treating pain -
XX
PS Claim 16; Fig 1; 185pp; English.
XX
CC The invention relates to Mrg (mas-related gene) protein, which is a
CC G-protein coupled receptor and drg-12 protein, which is a receptor. The
CC invention is useful for identifying compounds that bind to it, especially
CC agonists or antagonists. Administration of an agent (e.g. the identified
CC agonist) that increases the expression of Mrg in a mammal may be used for
CC treating impaired sensory perception in a mammal, especially pain. The
CC antagonist may also be useful for treating impaired sensory perception in
CC a mammal. The present sequence is human MrgX1 protein.
XX
SQ Sequence 322 AA;

```
  Query Match             100.0%;  Score 1661;  DB 23;  Length 322;
  Best Local Similarity   100.0%;  Pred. No. 4e-171;
  Matches  322;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MDPTISTLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDPTISTLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNA 60
```

```
Qy          61 FSIYILNLAAADFLFLSGRLIYSLLSFISIPHTISKILYPVMMFSYFAGLSFLSAVSTER 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 FSIYILNLAAADFLFLSGRLIYSLLSFISIPHTISKILYPVMMFSYFAGLSFLSAVSTER 120

Qy         121 CLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQTSDFIT 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 CLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQTSDFIT 180

Qy         181 VAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGIQFFLFL 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 VAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGIQFFLFL 240

Qy         241 WIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASE 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 WIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASE 300

Qy         301 VDEGGGQLPEEILELSGSRLEQ 322
               |||||||||||||||||||||||
Db         301 VDEGGGQLPEEILELSGSRLEQ 322
```

RESULT 5
AAE17074
ID AAE17074 standard; Protein; 322 AA.
XX
AC AAE17074;
XX
DT 18-APR-2002 (first entry)
XX
DE Human G-protein coupled receptor (GPCRx5) protein.
XX
KW Human; G-protein coupled receptor; GPCRx5; cerebroprotective; vomiting;
KW receptor-mediated disorder; therapy; urinary retention; allergy; obesity;
KW osteoporosis; angina pectoris; restenosis; atherosclerosis; hypotension;
KW anorexia; tumour; migraine; acute heart failure; ulcer; antiinflammatory;
KW stroke; hypertension; neuronal disorder; myocardial infarction psychotic;
KW depression; mental retardation; neurodegenerative disease; antibacterial;
KW Alzheimer's disease; dementia; ischaemia; Parkinson's disease; antiviral;
KW Huntington's disease; anxiety; antifungal; immunosuppressive; cytostatic;
KW vulnerary; analgesic; anorectic; anabolic; diuretic; cardiant; nootropic;
KW antiemetic; vasotropic; diabetes; cancer; tranquillizer; neuroleptic.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 28..52
FT /note= "Transmembrane domain"
FT Domain 60..88
FT /note= "Transmembrane domain"
FT Domain 97..116
FT /note= "Transmembrane domain"
FT Domain 140..160
FT /note= "Transmembrane domain"
FT Domain 178..200
FT /note= "Transmembrane domain"
FT Domain 255..275
FT /note= "Transmembrane domain"
XX
PN WO200198330-A2.
XX
PD 27-DEC-2001.
XX
PF 20-JUN-2001; 2001WO-BE00104.
XX
PR 20-JUN-2000; 2000US-212913P.
PR 11-JUL-2000; 2000US-217494P.
PR 26-JAN-2001; 2001EP-0870015.
PR 12-FEB-2001; 2001EP-0870024.
XX
PA (EURO-) EUROSCREEN SA.
XX
PI Lannoy V, Brezillon S, Detheux M, Parmentier M, Govarts C;
XX
DR WPI; 2002-130789/17.
DR N-PSDB; AAD27494.
XX
PT New G-protein coupled receptor, useful in the manufacture of
PT medicaments for treating receptor mediated disorders e.g. acute heart
PT failure and Alzheimer's disease -
XX
PS Disclosure; Page 23; 46pp; English.
XX
CC The present invention relates to a G-protein coupled receptor (GPCR) and
CC nucleotide encoding it. GPCR are useful in the manufacture of a
CC medicament for the prevention and/or treatment of receptor-mediated
CC disorders e.g. viral infections, virus and bacterial diseases, diseases
CC and disorders involving disturbances of cell migration, diseases or
CC perturbations of immune system including cancers, development of tumours
CC and tumour metastasis, inflammatory and neoplastic processes; bacterial
CC and fungal infections, in wound and bone healing, dysfunction of
CC regulatory growth functions; pains, diabetes, obesity, anorexia, bulimia,
CC urinary retention, osteoporosis, angina pectoris, atherosclerosis,
CC restenosis, diseases involving excessive or reduced proliferation or loss
CC of smooth muscle cells, aneurysm, stroke, ischaemia, ulcers, allergies,
CC benign prostatic hypertrophy, migraine, vomiting; blood circulating
CC affections including acute heart failure, hypotension, hypertension and
CC myocardial infarction psychotic; neuronal disorders such as anxiety,
CC schizophrenia, maniac depression, depression, delirium, dementia, severe
CC mental retardation; degenerative diseases; neurodegenerative diseases
CC such as Alzheimer's disease, Parkinson's disease; and dyskinesias e.g.
CC Huntington's disease or Gilles de la Tourette's syndrome and other
CC related diseases. The present sequence is GPCRx5 protein.
XX
SQ Sequence 322 AA;

Query Match 100.0%; Score 1661; DB 23; Length 322;
Best Local Similarity 100.0%; Pred. No. 4e-171;
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy           1 MDPTISTLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNA 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 MDPTISTLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNA 60

Qy          61 FSIYILNLAAADFLFLSGRLIYSLLSFISIPHTISKILYPVMMFSYFAGLSFLSAVSTER 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 FSIYILNLAAADFLFLSGRLIYSLLSFISIPHTISKILYPVMMFSYFAGLSFLSAVSTER 120

Qy         121 CLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQTSDFIT 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 CLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQTSDFIT 180

Qy         181 VAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGIQFFLFL 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 VAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGIQFFLFL 240

Qy         241 WIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASE 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 WIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASE 300

Qy         301 VDEGGGQLPEEILELSGSRLEQ 322
               |||||||||||||||||||||||
Db         301 VDEGGGQLPEEILELSGSRLEQ 322
```

RESULT 6
AAY30162
ID AAY30162 standard; protein; 322 AA.
XX
AC AAY30162;
XX
DT 26-NOV-1999 (first entry)
XX
DE Human dorsal root receptor 4 hDRR4.
XX
KW Dorsal root receptor; dorsal root ganglia; G-protein coupled receptor;
KW hDRR4; central nervous system; CNS; anaesthesia; analgesia; neuron; pain.

```
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 310
FT                   /note= "Encoded by GAC"
FT   Misc-difference 312
FT                   /note= "Encoded by ATC"
XX
PN   WO9932519-A1.
XX
PD   01-JUL-1999.
XX
PF   16-DEC-1998;   98WO-SE02348.
XX
PR   22-DEC-1997;   97SE-0004836.
XX
PA   (ASTR ) ASTRA AB.
PA   (ASTR-) ASTRA PHARMA INC.
XX
PI   Ahmad S,  Banville D,  Fortin Y,  Lembo P,  O'Donell D;
PI   Shen S;
XX
DR   WPI; 1999-405162/34.
DR   N-PSDB; AAZ10070.
XX
PT   Rat and human dorsal root receptors and related polynucleotides,
PT   useful for identifying agents for anaesthesia and analgesia
XX
PS   Claim 21; Page 52-54; 72pp; English.
XX
CC   This is the human dorsal root receptor 4 (hDRR4) protein sequence. This
CC   is a G protein coupled receptor that is expressed preferentially in
CC   dorsal root ganglia. hDRR4 can be used to create antibodies against
CC   hDRR4. The dorsal root ganglia area of the central nervous system (CNS)
CC   is densely innervated with primary or afferent neurons involved in
CC   transmission, modulation and sensation of pain. The DRR's which are
CC   expressed in this region of the CNS may be used for assays for the
CC   identification of new agents for anaesthesia and analgesia.
XX
SQ   Sequence   322 AA;

  Query Match            99.5%;  Score 1652;  DB 20;  Length 322;
  Best Local Similarity  99.1%;  Pred. No. 3.7e-170;
  Matches  319;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MDPTISTLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNA 60
              ||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDPTVSTLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNA 60

Qy         61 FSIYILNLAAADFLFLSGRLIYSLLSFISIPHTISKILYPVMMFSYFAGLSFLSAVSTER 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 FSIYILNLAAADFLFLSGRLIYSLLSFISIPHTISKILYPVMMFSYFAGLSFLSAVSTER 120

Qy        121 CLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQTSDFIT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQTSDFIT 180

Qy        181 VAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGIQFFLFL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGIQFFLFL 240

Qy        241 WIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 WIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASE 300

Qy        301 VDEGGGQLPEEILELSGSRLEQ 322
              |||||||||:| ||||||||||
Db        301 VDEGGGQLPQETLELSGSRLEQ 322


RESULT 7
AAY30161
ID   AAY30161 standard; protein; 322 AA.
XX
AC   AAY30161;
XX
DT   26-NOV-1999  (first entry)
XX
DE   Human dorsal root receptor 3 hDRR3.
XX
KW   Dorsal root receptor; dorsal root ganglia; G-protein coupled receptor;
KW   hDRR3; central nervous system; CNS; anaesthesia; analgesia; neuron; pain.
XX
OS   Homo sapiens.
XX
PN   WO9932519-A1.
XX
PD   01-JUL-1999.
XX
PF   16-DEC-1998;   98WO-SE02348.
XX
PR   22-DEC-1997;   97SE-0004836.
XX
PA   (ASTR ) ASTRA AB.
PA   (ASTR-) ASTRA PHARMA INC.
XX
PI   Ahmad S,  Banville D,  Fortin Y,  Lembo P,  O'Donell D;
PI   Shen S;
XX
DR   WPI; 1999-405162/34.
DR   N-PSDB; AAZ10069.
XX
PT   Rat and human dorsal root receptors and related polynucleotides,
PT   useful for identifying agents for anaesthesia and analgesia
XX
PS   Claim 16; Page 48-50; 72pp; English.
XX
CC   This is the human dorsal root receptor 3 (hDRR3) protein sequence. This
CC   is a G protein coupled receptor that is expressed preferentially in
CC   dorsal root ganglia. hDRR3 can be used to create antibodies against
CC   hDRR3. The dorsal root ganglia area of the central nervous system (CNS)
CC   is densely innervated with primary or afferent neurons involved in
CC   transmission, modulation and sensation of pain. The DRR's which are
CC   expressed in this region of the CNS may be used for assays for the
CC   identification of new agents for anaesthesia and analgesia.
XX
SQ   Sequence   322 AA;

  Query Match            97.8%;  Score 1624;  DB 20;  Length 322;
  Best Local Similarity  97.2%;  Pred. No. 4e-167;
  Matches  313;  Conservative    4;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 MDPTISTLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNA 60
              ||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDPTVSTLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNA 60

Qy         61 FSIYILNLAAADFLFLSGRLIYSLLSFISIPHTISKILYPVMMFSYFAGLSFLSAVSTER 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||:|
Db         61 FSIYILNLAAADFLFLSGRLIYSLLSFISIPHTISKILYPVMMFSYFAGLNFLSAVSTDR 120

Qy        121 CLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQTSDFIT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQTSDFIT 180

Qy        181 VAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGIQFFLFL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGIQFFLFL 240

Qy        241 WIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASE 300
              |||||||||||||||||||||||||||||||||||| |||||||||||||||||  |
Db        241 WIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSLRQRQNRQNLKLVLQRALQDTPE 300

Qy        301 VDEGGGQLPEEILELSGSRLEQ 322
              |||||| ||:| ||||||||||
Db        301 VDEGGGWLPQETLELSGSRLEQ 322
```

RESULT 8
AAB14846
ID AAB14846 standard; Protein; 322 AA.
XX
AC AAB14846;
XX
DT 19-DEC-2000 (first entry)
XX
DE Human novel G-protein coupled receptor #1.
XX
KW Human; novel G-protein coupled receptor; signal transduction;
KW disease diagnosis; drug screening; disease therapy.
XX
OS Homo sapiens.
XX
PN WO200040724-A1.
XX
PD 13-JUL-2000.
XX
PF 04-JAN-2000; 2000WO-US00052.
XX
PR 04-JAN-1999; 99US-0114666.
PR 14-JAN-1999; 99US-0115828.
XX
PA (LEXI-) LEXICON GENETICS INC.
XX
PI Nehls M, Wattler F;
XX
DR WPI; 2000-465986/40.
DR N-PSDB; AAA70342.
XX
PT New polynucleotides encoding novel G-protein coupled receptors useful
PT for diagnosis, drug screening, clinical trial monitoring and for the
PT treatment of physiological or behavioural disorders -
XX
PS Claim 1; Page 53-54; 61pp; English.
XX
CC The present sequence is the protein sequence for a novel human G-protein
CC coupled receptor (NGPCR). These proteins are involved in signal
CC transduction pathways in many cases. The protein contains seven
CC transmembrane domains, and is expressed in human testis, mammary gland
CC and salivary gland tissue. The protein, its gene, agonists, antagonists
CC and antibodies can be used to diagnose and treat diseases associated
CC with the inappropriate expression or expression of mutant versions of the
CC protein, for screening for drugs which can be used in the same manner,
CC and for elucidating the function of the protein.
XX
SQ Sequence 322 AA;

Query Match 83.7%; Score 1390; DB 21; Length 322;
Best Local Similarity 83.9%; Pred. No. 8.3e-142;
Matches 270; Conservative 16; Mismatches 36; Indels 0; Gaps 0;

```
Qy          1 MDPTISTLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNA 60
              || ||  | |||||||| ||| ||||||| | |||||||| ||||||||||||||||||||
Db          1 MDSTIPVLGTELTPINGREETPCYKQTLSFTGLTCIVSLVALTGNAVVLWLLGCRMRRNA 60

Qy         61 FSIYILNLAAADFLFLSGRLIYSLLSFISIPHTISKILYPVMMFSYFAGLSFLSAVSTER 120
              ||||||| ||||||||| :| | |  |:| | ||||| ||| | || ||| |||:||||
Db         61 VSIYILNLVAADFLFLSGHIIRSPLRLINIRHPISKILSPVMTFPYFIGLSMLSAISTER 120

Qy        121 CLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQTSDFIT 180
              |||:|||||| | || :||:|:|||||||||||||||| | |||||||| ||:||||||
Db        121 CLSILWPIWYHCRRPRYLSSVMCVLLWALSLLRSILEWMFCDFLFSGADSVWCETSDFIT 180

Qy        181 VAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGIQFFLFL 240
              :|||:|||||||||||||:|||||||:|||||||||||||||||||||||||||||||: ||
Db        181 IAWLVFLCVVLCGSSLVLLVRILCGSRKMPLTRLYVTILLTVLVFLLCGLPFGIQWALFS 240

Qy        241 WIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASE 300
              ||:| :|||||||||||||||||||||||||||||||||||||||||||||||||||| |
```

Db 241 RIHLDWKVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDTPE 300

Qy 301 VDEGGGQLPEEILELSGSRLEQ 322
||||||||:| ||||||||||
Db 301 VDEGGGQLPQETLELSGSRLEQ 322

RESULT 9
AAY90761
ID AAY90761 standard; Protein; 322 AA.
XX
AC AAY90761;
XX
DT 18-AUG-2000 (first entry)
XX
DE Human G protein-coupled receptor hHI7T213 SEQ ID NO:1.
XX
KW Human; G protein-coupled receptor; hippocampus; diagnosis; screening;
KW genetic disease; cellular function regulation.
XX
OS Homo sapiens.
XX
PN WO200020455-A1.
XX
PD 13-APR-2000.
XX
PF 30-SEP-1999; 99WO-JP05366.
XX
PR 01-OCT-1998; 98JP-0279535.
XX
PA (TAKE ) TAKEDA CHEM IND LTD.
XX
PI Watanabe T, Terao Y, Matsui H;
XX
DR WPI; 2000-303747/26.
DR N-PSDB; AAA29811.
XX
PT Human-derived G protein-coupled protein and encoding nucleic acid,
PT useful e.g. in determining ligands and treatment of diseases associated
PT with dysfunction of the protein -
XX
PS Claim 1; Page 90-91; 97pp; Japanese.
XX
CC The present sequence represents a human-derived G protein-coupled protein
CC designated hHI7T213, which is isolated from the human hippocampus.
CC The G protein-coupled receptor can be used for preventing, treating and
CC diagnosing genetic diseases associated with G protein-coupled protein,
CC and for regulating cellular functions. The protein can be used to
CC prevent and treat disorders associated with G protein-coupled protein
CC gene dysfunction. It can also be used to identify G protein-coupled
CC protein ligands and generating antibodies and antisera against the
CC protein. It is also useful in constructing recombinant receptor protein
CC expression systems, developing receptor-binding assay systems and
CC screening drug candidates, and can be used as a probe in the genetic
CC diagnosis of G protein-coupled protein disorders.
XX
SQ Sequence 322 AA;

Query Match 83.3%; Score 1383; DB 21; Length 322;
Best Local Similarity 83.5%; Pred. No. 4.7e-141;
Matches 269; Conservative 16; Mismatches 37; Indels 0; Gaps 0;

```
Qy          1 MDPTISTLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNA 60
              || ||  | |||||||| ||| ||||||| | |||||||| ||||||||||||||||||||
Db          1 MDSTIPVLGTELTPINGREETPCYKQTLSFTGLTCIVSLVALTGNAVVLWLLGCRMRRNA 60

Qy         61 FSIYILNLAAADFLFLSGRLIYSLLSFISIPHTISKILYPVMMFSYFAGLSFLSAVSTER 120
              ||||||| ||||||||| :| | |  |:| | ||||| ||| | || ||| |||:||||
Db         61 VSIYILNLVAADFLFLSGHIICSPLRLINIRHPISKILSPVMTFPYFIGLSMLSAISTER 120

Qy        121 CLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQTSDFIT 180
              |||:|||||| | || :||:|:|||||||||||||||| | |||||||| ||:||||||
Db        121 CLSILWPIWYHCRRPRYLSSVMCVLLWALSLLRSILEWMFCDFLFSGADSVWCETSDFIT 180
```

```
Qy        181 VAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGIQFFLFL 240
              :|||:|||||||||||||||:||||||||:|||||||||||||||||||||||||||||: ||
Db        181 IAWLVFLCVVLCGSSLVLLVRILCGSRKMPLTRLYVTILLTVLVFLLCGLPFGIQWALFS 240

Qy        241 WIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASE 300
               ||:| :|||||||||||||||||||||||||||||||||||||||||||||||||||  |
Db        241 RIHLDWKVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDTPE 300

Qy        301 VDEGGGQLPEEILELSGSRLEQ 322
              |||||| ||:| ||||||||||
Db        301 VDEGGGWLPQETLELSGSRLEQ 322


RESULT 10
AAE21296
ID   AAE21296 standard; Protein; 322 AA.
XX
AC   AAE21296;
XX
DT   01-JUL-2002  (first entry)
XX
DE   Human MrgX3 (mas-related gene) protein.
XX
KW   Human; mas-related gene; G-protein coupled receptor; drg-12 protein;
KW   receptor; sensory perception; pain; analgesic; MrgX3.
XX
OS   Homo sapiens.
XX
PN   WO200183555-A2.
XX
PD   08-NOV-2001.
XX
PF   04-MAY-2001; 2001WO-US14519.
XX
PR   04-MAY-2000; 2000US-202027P.
PR   01-AUG-2000; 2000US-222344P.
PR   03-NOV-2000; 2000US-0704707.
PR   19-APR-2001; 2001US-285493P.
XX
PA   (CALY ) CALIFORNIA INST OF TECHNOLOGY.
XX
PI   Anderson DJ,  Dong X,  Zylka M,  Han S,  Simon M;
XX
DR   WPI; 2002-171346/22.
DR   N-PSDB; AAD33751.
XX
PT   Isolated polypeptide, Mrg, which is a G-protein coupled receptor and an
PT   isolated polypeptide, drg-12, which is also a receptor, useful for
PT   identifying agonists or antagonists for treating pain  -
XX
PS   Claim 16; Page 130; 185pp; English.
XX
CC   The invention relates to Mrg (mas-related gene) protein, which is a
CC   G-protein coupled receptor and drg-12 protein, which is a receptor. The
CC   invention is useful for identifying compounds that bind to it, especially
CC   agonists or antagonists. Administration of an agent (e.g. the identified
CC   agonist) that increases the expression of Mrg in a mammal may be used for
CC   treating impaired sensory perception in a mammal, especially pain. The
CC   antagonist may also be useful for treating impaired sensory perception in
CC   a mammal. The present sequence is human MrgX3 protein.
XX
SQ   Sequence   322 AA;

  Query Match           83.3%;  Score 1383;  DB 23;  Length 322;
  Best Local Similarity   83.5%;  Pred. No. 4.7e-141;
  Matches  269;  Conservative   16;  Mismatches   37;  Indels    0;  Gaps    0;

Qy          1 MDPTISTLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNA 60
              || ||  | |||||||| ||| ||||||| | |||||||| ||||||||||||||||||||
Db          1 MDSTIPVLGTELTPINGREETPCYKQTLSFTGLTCIVSLVALTGNAVVLWLLGCRMRRNA 60

Qy         61 FSIYILNLAAADFLFLSGRLIYSLLSFISIPHTISKILYPVMMFSYFAGLSFLSAVSTER 120
              ||||||| ||||||||| :| | |  |:| | ||||| ||| | || ||| |||:||||
Db         61 VSIYILNLVAADFLFLSGHIICSPLRLINIRHPISKILSPVMTFPYFIGLSMLSAISTER 120

Qy        121 CLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQTSDFIT 180
              |||:|||||  | || :||:|:|||||||||||||||| | |||||||| ||:||||||
Db        121 CLSILWPIWYHCRRPRYLSSVMCVLLWALSLLRSILEWMFCDFLFSGADSVWCETSDFIT 180

Qy        181 VAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGIQFFLFL 240
              :|||:|||||||||||||||:||||||||:|||||||||||||||||||||||||||||: ||
Db        181 IAWLVFLCVVLCGSSLVLLVRILCGSRKMPLTRLYVTILLTVLVFLLCGLPFGIQWALFS 240

Qy        241 WIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASE 300
               ||:| :|||||||||||||||||||||||||||||||||||||||||||||||||||  |
Db        241 RIHLDWKVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDTPE 300

Qy        301 VDEGGGQLPEEILELSGSRLEQ 322
              |||||| ||:| ||||||||||
Db        301 VDEGGGWLPQETLELSGSRLEQ 322


RESULT 11
AAU97598
ID   AAU97598 standard; Protein; 337 AA.
XX
AC   AAU97598;
XX
DT   12-AUG-2002  (first entry)
XX
DE   Human G-protein coupled receptor (GPCR).
XX
KW   Human; G-protein coupled; receptor; GPCR; human protease;
KW   human therapeutic protein; query sequence; search; chromosome 3;
KW   sequence database; non-human transgenic animal; gene therapy.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Domain          41..61
FT                   /label= Helix_1
FT   Domain          75..95
FT                   /label= Helix_2
FT   Domain          112..132
FT                   /label= Helix_3
FT   Modified-site   133..135
FT                   /note= "Protein kinase C (PKC) phosphorylation site"
FT   Domain          151..171
FT                   /label= Helix_4
FT   Modified-site   169..172
FT                   /note= "Casein kinase II (CK2) phosphorylation site"
FT   Modified-site   181..184
FT                   /note= "Casein kinase II (CK2) phosphorylation site"
FT   Domain          193..213
FT                   /label= Helix_5
FT   Modified-site   221..223
FT                   /note= "Protein kinase C (PKC) phosphorylation site"
FT   Domain          229..249
FT                   /label= Helix_6
FT   Modified-site   244..249
FT                   /note= "Myristoylation site"
FT   Modified-site   248..253
FT                   /note= "Myristoylation site"
FT   Domain          261..281
FT                   /label= Helix_7
FT   Modified-site   279..282
FT                   /note= "Asn glycosylation site"
FT   Modified-site   292..294
FT                   /note= "Protein kinase C (PKC) phosphorylation site"
FT   Modified-site   333.336
FT                   /note= "Casein kinase II (CK2) phosphorylation site"
XX
PN   WO200234914-A1.
XX
PD   02-MAY-2002.
```

XX
PF 10-OCT-2001; 2001WO-US31592.
XX
PR 25-OCT-2000; 2000US-0695045.
PR 31-MAY-2001; 2001US-0867570.
XX
PA (PEKE ) PE CORP NY.
XX
PI Wei M, Zhao Q, Woodage T, Di Francesco V, Beasley EM;
XX
DR WPI; 2002-463360/49.
DR N-PSDB; ABK52822, ABK52823.
XX
PT Novel isolated G-protein coupled receptor peptide useful for treating
PT disorder characterised by absence of inappropriate or unwanted
PT expression of the receptor protein, and as immunogens to raise
PT antibodies -
XX
PS Claim 1; Fig 2; 75pp; English.
XX
CC The present invention relates to a new G-protein coupled receptor
CC (GPCR) peptide. The invention is useful for identifying a modulator
CC of GPCR and for treating a disease or condition mediated by a human
CC protease. The invention is also useful as models for the development
CC of human therapeutics, for identifying therapeutic proteins, as
CC targets for development of human therapeutic agents, and as query
CC sequence to perform a search against sequence databases to, for
CC e.g., identify other family members of related sequences. The
CC vector of the invention is useful for producing a GPCR protein or
CC peptide, for conducting cell-based assays involving the GPCR protein
CC or its fragment, for identifying GPCR protein mutants whose
CC functions are affected, and to produce non-human transgenic animals.
CC The present amino acid sequence represents the human G-protein
CC coupled receptor (GPCR) protein of the invention. This sequence is
CC encoded by the human G-protein coupled receptor (GPCR) gene located
CC on chromosome 3.
XX
SQ Sequence 337 AA;

Query Match 83.3%; Score 1383; DB 23; Length 337;
Best Local Similarity 83.5%; Pred. No. 5e-141;
Matches 269; Conservative 16; Mismatches 37; Indels 0; Gaps 0;

```
Qy          1 MDPTISTLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNA 60
              || ||  | ||||||| ||| ||||||| | |||||||| ||||||||||||||||||||
Db         16 MDSTIPVLGTELTPINGREETPCYKQTLSFTGLTCIVSLVALTGNAVVLWLLGCRMRRNA 75

Qy         61 FSIYILNLAAADFLFLSGRLIYSLLSFISIPHTISKILYPVMMFSYFAGLSFLSAVSTER 120
              ||||||| ||||||||| :| | |  |:| | ||||| ||| | || ||| |||:||||
Db         76 VSIYILNLVAADFLFLSGHIICSPLRLINIRHPISKILSPVMTFPYFIGLSMLSAISTER 135

Qy        121 CLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQTSDFIT 180
              |||:|||||| | || :||:|:|||||||||||||||| | |||||||| ||:||||||
Db        136 CLSILWPIWYHCRRPRYLSSVMCVLLWALSLLRSILEWMFCDFLFSGADSVWCETSDFIT 195

Qy        181 VAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGIQFFLFL 240
              :|||:|||||||||||||:|||||||:|||||||||||||||||||||||||||||: ||
Db        196 IAWLVFLCVVLCGSSLVLLVRILCGSRKMPLTRLYVTILLTVLVFLLCGLPFGIQWALFS 255

Qy        241 WIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASE 300
              ||:| :|||||||||||||||||||||||||||||||||||||||||||||||||||  |
Db        256 RIHLDWKVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDTPE 315

Qy        301 VDEGGGQLPEEILELSGSRLEQ 322
              |||||| ||:| ||||||||||
Db        316 VDEGGGWLPQETLELSGSRLEQ 337
```

RESULT 12
AAY90762
ID AAY90762 standard; Protein; 322 AA.
XX
AC AAY90762;

XX
DT 18-AUG-2000 (first entry)
XX
DE Human G protein-coupled receptor hHI7T213V SEQ ID NO:2.
XX
KW Human; G protein-coupled receptor; hippocampus; diagnosis; screening;
KW genetic disease; cellular function regulation.
XX
OS Homo sapiens.
XX
PN WO200020455-A1.
XX
PD 13-APR-2000.
XX
PF 30-SEP-1999; 99WO-JP05366.
XX
PR 01-OCT-1998; 98JP-0279535.
XX
PA (TAKE ) TAKEDA CHEM IND LTD.
XX
PI Watanabe T, Terao Y, Matsui H;
XX
DR WPI; 2000-303747/26.
DR N-PSDB; AAA29812.
XX
PT Human-derived G protein-coupled protein and encoding nucleic acid,
PT useful e.g. in determining ligands and treatment of diseases associated
PT with dysfunction of the protein -
XX
PS Claim 2; Page 92-93; 97pp; Japanese.
XX
CC The present sequence represents a human-derived G protein-coupled protein
CC designated hHI7T213V, which is isolated from the human hippocampus.
CC The G protein-coupled receptor can be used for preventing, treating and
CC diagnosing genetic diseases associated with G protein-coupled protein,
CC and for regulating cellular functions. The protein can be used to
CC prevent and treat disorders associated with G protein-coupled protein
CC gene dysfunction. It can also be used to identify G protein-coupled
CC protein ligands and generating antibodies and antisera against the
CC protein. It is also useful in constructing recombinant receptor protein
CC expression systems, developing receptor-binding assay systems and
CC screening drug candidates, and can be used as a probe in the genetic
CC diagnosis of G protein-coupled protein disorders.
XX
SQ Sequence 322 AA;

Query Match 83.0%; Score 1378; DB 21; Length 322;
Best Local Similarity 83.2%; Pred. No. 1.6e-140;
Matches 268; Conservative 17; Mismatches 37; Indels 0; Gaps 0;

```
Qy          1 MDPTISTLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNA 60
              || ||  | ||||||| ||| ||||||| | |||||||| ||||||||||||||||||||
Db          1 MDSTIPVLGTELTPINGREETPCYKQTLSFTGLTCIVSLVALTGNAVVLWLLGCRMRRNA 60

Qy         61 FSIYILNLAAADFLFLSGRLIYSLLSFISIPHTISKILYPVMMFSYFAGLSFLSAVSTER 120
              ||||||| ||||||||| :| | |  |:| | ||||| ||| | || ||| |||:||||
Db         61 VSIYILNLVAADFLFLSGHIICSPLRLINIRHPISKILSPVMTFPYFIGLSMLSAISTER 120

Qy        121 CLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQTSDFIT 180
              |||:|||||| | || :||:|:|||||||||||||||| | |||||||| ||:||||||
Db        121 CLSILWPIWYHCRRPRYLSSVMCVLLWALSLLRSILEWMFCDFLFSGANSVWCETSDFIT 180

Qy        181 VAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGIQFFLFL 240
              :|||:|||||||||||||:|||||||:|||||||||||||||||||||||||||||: ||
Db        181 IAWLVFLCVVLCGSSLVLLVRILCGSRKMPLTRLYVTILLTVLVFLLCGLPFGIQWALFS 240

Qy        241 WIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASE 300
              ||:| :|||||||||||||||||||||||||||||||||||||||||||||||||||  |
Db        241 RIHLDWKVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDTPE 300

Qy        301 VDEGGGQLPEEILELSGSRLEQ 322
              |||||| ||:| ||||||||||
Db        301 VDEGGGWLPQETLELSGSRLEQ 322
```

RESULT 13
AAY30159
ID AAY30159 standard; protein; 322 AA.
XX
AC AAY30159;
XX
DT 26-NOV-1999 (first entry)
XX
DE Human dorsal root receptor 1 hDRR1.
XX
KW Dorsal root receptor; dorsal root ganglia; G-protein coupled receptor;
KW hDRR1; central nervous system; CNS; anaesthesia; analgesia; neuron; pain.
XX
OS Homo sapiens.
XX
PN WO9932519-A1.
XX
PD 01-JUL-1999.
XX
PF 16-DEC-1998; 98WO-SE02348.
XX
PR 22-DEC-1997; 97SE-0004836.
XX
PA (ASTR ) ASTRA AB.
PA (ASTR-) ASTRA PHARMA INC.
XX
PI Ahmad S, Banville D, Fortin Y, Lembo P, O'Donell D;
PI Shen S;
XX
DR WPI; 1999-405162/34.
DR N-PSDB; AAZ10067.
XX
PT Rat and human dorsal root receptors and related polynucleotides,
PT useful for identifying agents for anaesthesia and analgesia
XX
PS Claim 6; Page 39-41; 72pp; English.
XX
CC This is the human dorsal root receptor 1 (hDRR1) protein sequence. This
CC is a G protein coupled receptor that is expressed preferentially in
CC dorsal root ganglia. hDRR1 can be used to create antibodies against
CC hDRR1. The dorsal root ganglia area of the central nervous system (CNS)
CC is densely innervated with primary or afferent neurons involved in
CC transmission, modulation and sensation of pain. The DRR's which are
CC expressed in this region of the CNS may be used for assays for the
CC identification of new agents for anaesthesia and analgesia.
XX
SQ Sequence 322 AA;

Query Match 82.7%; Score 1373; DB 20; Length 322;
Best Local Similarity 82.0%; Pred. No. 5.7e-140;
Matches 264; Conservative 22; Mismatches 36; Indels 0; Gaps 0;

```
Qy          1 MDPTISTLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNA 60
              |||||  | |:|||||| ||| || |||| | ||||:||| ||||||||||||||||||
Db          1 MDPTIPVLGTKLTPINGREETPCYNQTLSFTGLTCIISLVALTGNAVVLWLLGCRMRRNA 60

Qy         61 FSIYILNLAAADFLFLSGRLIYSLLSFISIPHTISKILYPVMMFSYFAGLSFLSAVSTER 120
              ||||||| ||:|||||| :|:| |  |:| | ||||| ||| | || ||| |||:||||
Db         61 VSIYILNLVAANFLFLSGHIIFSPLPLINIRHPISKILSPVMTFPYFIGLSMLSAISTER 120

Qy        121 CLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQTSDFIT 180
              |||:|||||| | || :||:|:|||||||||||||| | ||||||:| ||:||||||
Db        121 CLSILWPIWYHCRRPRYLSSVMCVLLWALSLLRSILEWMFCDFLFSGANSVWCETSDFIT 180

Qy        181 VAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGIQFFLFL 240
              :|||:||||||||||||:||||||||:|||||||||||||||||||||||: ||
Db        181 IAWLVFLCVVLCGSSLVLLVRILCGSRKMPLTRLYVTILLTVLVFLLCGLPFGIQWALFS 240

Qy        241 WIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASE 300
              ||:| :|||||||||||||||||||||||||||||||||||||||||||||||||   |
Db        241 RIHLDWKVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDTPE 300

Qy        301 VDEGGGQLPEEILELSGSRLEQ 322
              |||||| ||:| |||||||:|||
Db        301 VDEGGGWLPQETLELSGSKLEQ 322
```

RESULT 14
AAY87664
ID AAY87664 standard; Protein; 322 AA.
XX
AC AAY87664;
XX
DT 11-AUG-2000 (first entry)
XX
DE Human G protein-coupled receptor protein #2.
XX
KW G protein-coupled receptor protein; human; treatment; diagnosis;
KW genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200020456-A1.
XX
PD 13-APR-2000.
XX
PF 30-SEP-1999; 99WO-JP05365.
XX
PR 01-OCT-1998; 98JP-0279512.
PR 20-OCT-1998; 98JP-0298667.
XX
PA (TAKE ) TAKEDA CHEM IND LTD.
XX
PI Watanabe T, Terao Y, Fukusumi S;
XX
DR WPI; 2000-303748/26.
DR N-PSDB; AAA12208.
XX
PT Novel human-originated G protein-coupled protein and encoding nucleic
PT acid, useful for e.g. identifying ligands and treating diseases
PT associated with dysfunction of the protein -
XX
PS Claim 2; Page 97-98; 105pp; Japanese.
XX
CC This invention describes a novel human G protein-coupled receptor
CC protein (I). The products of the invention can be used for preventing,
CC treating and diagnosing gene diseases and for regulating cellular
CC functions. The protein can be used to identify ligands and generate
CC antibodies and antisera. It is also useful in the construction of a
CC recombinant receptor protein expression system, developing a
CC receptor-binding assay system and screening drug candidates and as a
CC probe in diagnosing genetic disorders involving G protein-coupled
CC protein. This sequence represents a human G-protein coupled receptor
CC protein described in the method of the invention.
XX
SQ Sequence 322 AA;

Query Match 79.1%; Score 1314; DB 21; Length 322;
Best Local Similarity 80.6%; Pred. No. 1.4e-133;
Matches 258; Conservative 20; Mismatches 42; Indels 0; Gaps 0;

```
Qy          1 MDPTISTLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNA 60
              ||||:  | |:|||||| ||| ||||||| ||||||:||||||||||||||||||||||
Db          1 MDPTVPVLGTKLTPINGREETPCYKQTLSFTVLTCIISLVGLTGNAVVLWLLGCRMRRNA 60

Qy         61 FSIYILNLAAADFLFLSGRLIYSLLSFISIPHTISKILYPVMMFSYFAGLSFLSAVSTER 120
              |||||||||||||||| ::| | |  |:| | | |||  || | || ||| |||:||||
Db         61 VSIYILNLAAADFLFLSFQIIRSPLRLINISHLIRKILVSVMTFPYFTGLSMLSAISTER 120

Qy        121 CLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQTSDFIT 180
              |||||||||||| |||||||||||| |||| |:|||  | |||||||:||:|||||
Db        121 CLSVLWPIWYRCRRPTHLSAVVCVLLWGLSLLFSMLEWRFCDFLFSGADSSWCETSDFIP 180

Qy        181 VAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGIQFFLFL 240
```

```
                ||||||||||||| ||||||:||||||||:||||||||||||||||||||||||||||    |
Db        181 VAWLIFLCVVLCVSSLVLLVRILCGSRKMPLTRLYVTILLTVLVFLLCGLPFGILGALIY 240

Qy        241 WIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASE 300
               :|:: |||:|||:|| : ||:|||||||||||||||||||||||||||||||||||||  |
Db        241 RMHLNLEVLYCHVYLVCMSLSSLNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDKPE 300

Qy        301 VDEGGGQLPEEILELSGSRL 320
               ||:| |||||| ||||||||
Db        301 VDKGEGQLPEESLELSGSRL 320
```

RESULT 15
AAY30160
ID  AAY30160 standard; protein; 322 AA.
XX
AC  AAY30160;
XX
DT  26-NOV-1999  (first entry)
XX
DE  Human dorsal root receptor 2 hDRR2.
XX
KW  Dorsal root receptor; dorsal root ganglia; G-protein coupled receptor;
KW  hDRR2; central nervous system; CNS; anaesthesia; analgesia; neuron; pain.
XX
OS  Homo sapiens.
XX
PN  WO9932519-A1.
XX
PD  01-JUL-1999.
XX
PF  16-DEC-1998;  98WO-SE02348.
XX
PR  22-DEC-1997;  97SE-0004836.
XX
PA  (ASTR ) ASTRA AB.
PA  (ASTR-) ASTRA PHARMA INC.
XX
PI  Ahmad S,  Banville D,  Fortin Y,  Lembo P,  O'Donell D;
PI  Shen S;
XX
DR  WPI; 1999-405162/34.
DR  N-PSDB; AAZ10068.
XX
PT  Rat and human dorsal root receptors and related polynucleotides,
PT  useful for identifying agents for anaesthesia and analgesia
XX
PS  Claim 11; Page 43-45; 72pp; English.
XX
CC  This is the human dorsal root receptor 2 (hDRR2) protein sequence. This
CC  is a G protein coupled receptor that is expressed preferentially in
CC  dorsal root ganglia. hDRR2 can be used to create antibodies against
CC  hDRR2. The dorsal root ganglia area of the central nervous system (CNS)
CC  is densely innervated with primary or afferent neurons involved in
CC  transmission, modulation and sensation of pain. The DRR's which are
CC  expressed in this region of the CNS may be used for assays for the
CC  identification of new agents for anaesthesia and analgesia.
XX
SQ  Sequence  322 AA;

  Query Match           79.0%;  Score 1313;  DB 20;  Length 322;
  Best Local Similarity  80.1%;  Pred. No. 1.8e-133;
  Matches  257;  Conservative   21;  Mismatches   43;  Indels    0;  Gaps    0;

```
Qy          1 MDPTISTLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNA 60
              ||||:  | |||||||| ||| ||||||| | |||||||| ||||||||||||||||||
Db          1 MDPTVPVLGTELTPINGREETPCYKQTLSFTGLTCIVSLVALTGNAVVLWLLGCRMRRNA 60

Qy         61 FSIYILNLAAADFLFLSGRLIYSLLSFISIPHTISKILYPVMMFSYFAGLSFLSAVSTER 120
               |||||||| ||||||||| :| | |  |:| | ||||| ||| | || ||| |:|:||||
Db         61 VSIYILNLVAADFLFLSGHIICSPLRLINISHPISKILSPVMTFPYFIGLSMLNAISTER 120

Qy        121 CLSVLWPIWYRCHRPTHLSAVVCVLLWALSLLRSILEWMLCGFLFSGADSAWCQTSDFIT 180
               |||:|||||| | || :||:|:|||||| |||||||||| | ||||||||  |:||||||
Db        121 CLSILWPIWYHCRRPRYLSSVMCVLLWAPSLLRSILEWMFCDFLFSGADSVRCETSDFIT 180

Qy        181 VAWLIFLCVVLCGSSLVLLIRILCGSRKIPLTRLYVTILLTVLVFLLCGLPFGIQFFLFL 240
               :|||:|| |||||||||:||||||||:||||||||||||||||||||||||||||: ||
Db        181 IAWLVFLRVVLCGSSLVLLVRILCGSRKMPLTRLYVTILLTVLVFLLCGLPFGIQWALFS 240

Qy        241 WIHVDREVLFCHVHLVSIFLSALNSSANPIIYFFVGSFRQRQNRQNLKLVLQRALQDASE 300
               ||:| :|||||||||||||||||||||||||||:||||| |||: ||||||| |||  |
Db        241 RIHLDWKVLFCHVHLVSIFLSALNSSANPIIYFFMGSFRQLQNRKTLKLVLQRDLQDTPE 300

Qy        301 VDEGGGQLPEEILELSGSRLE 321
               |||||  ||:| ||||||:||
Db        301 VDEGGWWLPQETLELSGSKLE 321
```

Search completed: July  2, 2003, 19:20:44
Job time : 74 secs